



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 94257**

**TO: Phillip Gambel**  
**Location: CM1/9E12/8B03**  
**Art Unit: 1644**  
**Wednesday, May 28, 2003**

**Case Serial Number: 919408**

**From: Susan Hanley**  
**Location: Biotech-Chem Library**  
**CM1-6B05**  
**Phone: 305-4053**

**[susan.hanley@uspto.gov](mailto:susan.hanley@uspto.gov)**

### **Search Notes**

94257

**From:** STIC-ILL  
**Sent:** Monday, May 19, 2003 8:32 AM  
**To:** STIC-Biotech/ChemLib  
**Subject:** FW: 09/919408 lemiskchka amd

-----Original Message-----

**From:** Gambel, Phillip  
**Sent:** Monday, May 19, 2003 8:29 AM  
**T :** STIC-ILL  
**Subject:** 09/919408 lemiskchka amd

stic

please perform a sequence and a sequence interference search for

SEQ ID NO:1

SEQ ID NO: 2

SEQ ID NO:3

SEQ ID NO:4

hard copy

thanx

phillip gambel  
art unit 1644  
308-3997

1644 mailbox 9e12

CRFE

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

## TYPE OF SEARCH:

NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

## VENDOR/COST (where applic.)

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 02:59:08 ; Search time 8656.83 Seconds  
(without alignments)  
11608.404 Million cell updates/sec

Title: US-09-919-408-1  
Perfect score: 3453  
Sequence: 1 GCGCCCTGGTACCGCGGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.mu.\*
- 20: em.om.\*
- 21: em.or.\*
- 22: em.ov.\*
- 23: em.pat.\*
- 24: em.ph.\*
- 25: em.pl.\*
- 26: em.ro.\*
- 27: em.sts.\*
- 28: em.un.\*
- 29: em.vi.\*
- 30: em.htg.hum.\*
- 31: em.htg.inv.\*
- 32: em.htg.Other.\*
- 33: em.htg.mus.\*
- 34: em.htg.pln.\*
- 35: em.htg.rod.\*
- 36: em.htg.mam.\*
- 37: em.htg.vrt.\*
- 38: em.sy.\*
- 39: em.htgo.hum.\*
- 40: em.htgo.mus.\*
- 41: em.htgo.Other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3453	100.0	3453	6	AR005211	AR005211 Sequence
2	3453	100.0	3453	6	AR071703	AR071703 Sequence
3	3453	100.0	3453	6	I25169	I25169 Sequence 1
4	3453	100.0	3453	6	I40600	I40600 Sequence 1
5	3426	99.2	3426	10	MUSFLK2	M64689 Mouse flk-2
6	3344	96.8	3521	6	I44732	I44732 Sequence 1
7	3344	96.8	3521	10	MMFLT3	X59398 Mouse Flt3
8	2227.2	64.5	3501	6	AR005212	AR005212 Sequence
9	2227.2	64.5	3501	6	AR071704	AR071704 Sequence
10	2227.2	64.5	3501	6	I25170	I25170 Sequence 3
11	2227.2	64.5	3501	6	I40601	I40601 Sequence 3
12	2204.8	63.9	3476	6	I40109	I40109 Sequence 1
13	2203	63.8	3475	6	I44733	I44733 Sequence 3
14	2203	63.8	3475	9	HS002687	U02687 Human growt
15	2191.4	63.5	3422	9	HSFLT3RTK	Z26652 H.sapiens F
16	2081.8	60.3	3120	6	I44518	I44518 Sequence 22
17	1206	34.9	1894	6	ARI49571	ARI49571 Sequence
18	474.6	13.7	181003	2	AC127549	AC127549 Mus muscu
19	427.4	12.4	429	10	MUS3RTK	L36163 Mus musculu
20	363.2	10.5	480	9	HOM3RTK	L36162 Homo sapien
21	357.4	10.4	3992	6	AX394288	AX394288 Sequence
22	357.4	10.4	3992	9	HSCFMS	X03663 Human mRNA
23	349.8	10.1	3545	5	AF153446	AF153446 Danio rer
24	344.2	10.0	389	10	AY094358	AY094358 Rattus no
25	341.8	9.9	4577	14	FCSSMONC	K01643 Feline sarc
26	340.2	9.9	3828	4	CATFMSC	J03149 Cat (F.dome
27	336.8	9.8	4222	4	S76596	S76596 c-kit=recep
28	335.6	9.7	3705	10	RRCKTTRTK	X62491 R.rattus mr
29	332.6	9.6	2176	4	AF263826	AF263826 Bos tauru
30	332.6	9.6	2176	4	AF263827	AF263827 Bos tauru
31	332.6	9.6	3069	4	BOVCKR	D16680 Bovine mRNA
32	331.4	9.6	2919	6	A82459	A82459 Sequence 47
33	331.4	9.6	2960	4	SSAJ3228	AJ223228 Sus scrof
34	331.4	9.6	2960	4	SSAJ3229	AJ223229 Sus scrof
35	331.4	9.6	2960	4	SSAJ3230	AJ223230 Sus scrof
36	331.2	9.6	3816	10	RATCKITPO	D12524 Rattus norv
37	330.4	9.6	3679	10	RRCSEF1	X61479 Rat mRNA fo
38	330.2	9.6	2894	4	ECA224643	AJ224643 Equus cab
39	330.2	9.6	2894	4	ECA224644	AJ224644 Equus cab
40	330.2	9.6	2973	4	ECA224642	AJ224642 Equus cab
41	330	9.6	3154	4	AF044249	AF044249 Canis fam
42	329.8	9.6	2979	4	AF448148	AF448148 Canis fam
43	329.2	9.5	2894	4	ECA224645	AJ224645 Equus cab
44	329	9.5	5084	6	AX195908	AX195908 Sequence
45	329	9.5	5084	6	AX331941	AX331941 Sequence

ALIGNMENTS

RESULT 1	AR005211	AR005211	Sequence 1	from patent US 5747651.	3453 bp	DNA	linear	PAT 04-DEC-1998
LOCUS	AR005211	Sequence 1	from patent US 5747651.					
DEFINITION	AR005211	Sequence 1	from patent US 5747651.					
ACCESSION	AR005211	Sequence 1	from patent US 5747651.					
VERSION	AR005211.1	GI:3966090						
KEYWORDS								
SOURCE	Unknown.							
ORGANISM	Unknown.							
REFERENCE	1 (bases 1 to 3453)							
AUTHORS	Lemischka, I.R.							
TITLE	Antibodies against tyrosine kinase receptor flk-1							
JOURNAL	Patent: US 5747651-A 1 05-MAY-1998;							
FEATURES	Location/Qualifiers							





Db 361 AGCTCCCTGGGCTGCCAGCGCACTTTGATTTTCAAAAACAGAGGAATCGTTTCCATGGCC 420  
Qy 421 ATCTTGAACGTGACAGAGACCCAGGAGGAATACCTACTCTCCATATATCAGAGCAAGC 480  
Db 421 ATCTTGAACGTGACAGAGACCCAGGAGGAATACCTACTCTCCATATATCAGAGCAAGC 480  
Qy 481 GCCAACTACACAGTACTGTTTCAACAGTGAATGTAAGAGATACACAGCTGTATGCTGAAG 540  
Db 481 GCCAACTACACAGTACTGTTTCAACAGTGAATGTAAGAGATACACAGCTGTATGCTGAAG 540  
Qy 541 AGACCTTACTTTAGGAAGATGGAACACAGATGCATCTCTGCATCTCCGAGGGTGT 600  
Db 541 AGACCTTACTTTAGGAAGATGGAACACAGATGCATCTCTGCATCTCCGAGGGTGT 600  
Qy 601 CCGAGAGCCACTGAGGAGTGGTCTCTGCAGCTCCACAGGGAAGCTGTAAAGAGAA 660  
Db 601 CCGAGAGCCACTGAGGAGTGGTCTCTGCAGCTCCACAGGGAAGCTGTAAAGAGAA 660  
Qy 661 GGCCCTGCTGTGTCAGAAAGGAGAAAGGTACTTTCATGAGTTGTTCCGGAACAGATC 720  
Db 661 GGCCCTGCTGTGTCAGAAAGGAGAAAGGTACTTTCATGAGTTGTTCCGGAACAGATC 720  
Qy 721 AGATGCTGTGTAAGAAATGCACTGGCGGGAATGCAACAGCTGTTCACCATAGACTA 780  
Db 721 AGATGCTGTGTAAGAAATGCACTGGCGGGAATGCAACAGCTGTTCACCATAGACTA 780  
Qy 781 AACGAGGCTCTCAGACACACTCCCGCAGTTATCTGAAAGTGGGGAACCTGTGG 840  
Db 781 AACGAGGCTCTCAGACACACTCCCGCAGTTATCTGAAAGTGGGGAACCTGTGG 840  
Qy 841 ATCAGGTTGTAAGGCACTCCATGTGAACCATGGATTTCGGCTCACCTGGAGCTGSAAGAC 900  
Db 841 ATCAGGTTGTAAGGCACTCCATGTGAACCATGGATTTCGGCTCACCTGGAGCTGSAAGAC 900  
Qy 901 AAAGCCCTGGAGGAGGCACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 960  
Db 901 AAAGCCCTGGAGGAGGCACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 960  
Qy 961 ATTCGGATTCCTTGGGCTTTGCTCTCCGTTGGGAAGAACACACCGGATATTACACC 1020  
Db 961 ATTCGGATTCCTTGGGCTTTGCTCTCCGTTGGGAAGAACACACCGGATATTACACC 1020  
Qy 1021 TGCTTCTTCTCAAAAGCACCCAGCAGTACGCTTGGTGGTACCATCTAGAAAAAGGGTTT 1080  
Db 1021 TGCTTCTTCTCAAAAGCACCCAGCAGTACGCTTGGTGGTACCATCTAGAAAAAGGGTTT 1080  
Qy 1081 ATAAACGCTACCACTGCAAGAGAGATGTAATTTGACCGGTAGAAAAAGTTCTGCTTC 1140  
Db 1081 ATAAACGCTACCACTGCAAGAGAGATGTAATTTGACCGGTAGAAAAAGTTCTGCTTC 1140  
Qy 1141 TCAGTCAAGTTTAAAGCGTACCCACGAATCCGATGCGAGTGATCTTCTCAAGCCCTCA 1200  
Db 1141 TCAGTCAAGTTTAAAGCGTACCCACGAATCCGATGCGAGTGATCTTCTCAAGCCCTCA 1200  
Qy 1201 TTTCTCTTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
Db 1201 TTTCTCTTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
Qy 1261 AAGAACAGCCAGGAGAGTACATATTCATGCAAGAAATGATGACGCCAGTTCACCAA 1320  
Db 1261 AAGAACAGCCAGGAGAGTACATATTCATGCAAGAAATGATGACGCCAGTTCACCAA 1320  
Qy 1321 ATGTTCAAGCTGATATTAAGAAAGAAACCTCAAGTCTAGCAATGCCTCAGCCAGCCAG 1380  
Db 1321 ATGTTCAAGCTGATATTAAGAAAGAAACCTCAAGTCTAGCAATGCCTCAGCCAGCCAG 1380  
Qy 1381 GCGTCTTCTCTGATGGCTACCCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAC 1440  
Db 1381 GCGTCTTCTCTGATGGCTACCCGCTACCTCTTGGACCTGGAAGAGTGTTCGGAC 1440  
Qy 1441 AAATCTCCAAATTCACAGGAGAAATCCAGAAAGAGTTTGAATATAAAGAGCTAACAGA 1500  
Db 1441 AAATCTCCAAATTCACAGGAGAAATCCAGAAAGAGTTTGAATATAAAGAGCTAACAGA 1500

Qy 1501 AAAGTGTGTGCCAGTGGTGTGAGCAGTACTCTAAATATGAGTGAGGCCGGGAAGG 1560  
Db 1501 AAAGTGTGTGCCAGTGGTGTGAGCAGTACTCTAAATATGAGTGAGGCCGGGAAGG 1560  
Qy 1561 CTCTGTGTCAAATGCTGTGCGTACAAATCTATGGCAGCTCTTGGAAACCATCTTTT 1620  
Db 1561 CTCTGTGTCAAATGCTGTGCGTACAAATCTATGGCAGCTCTTGGAAACCATCTTTT 1620  
Qy 1621 AACTCACCAGGCCCTTCCCTTTCATCAAGACACATCTCTCTATCGACCATTTGG 1680  
Db 1621 AACTCACCAGGCCCTTCCCTTTCATCAAGACACATCTCTCTATCGACCATTTGG 1680  
Qy 1681 CTCTGTCTCCCTTTCATTTGTTGTTGATCTGCCCAAAATACAAAAAGCAA 1740  
Db 1681 CTCTGTCTCCCTTTCATTTGTTGTTGATCTGCCCAAAATACAAAAAGCAA 1740  
Qy 1741 TTTAGGTACGAGAGTACGTGCAGATGATCCAGGTGACTGGCCCTCGATACAGTAC 1800  
Db 1741 TTTAGGTACGAGAGTACGTGCAGATGATCCAGGTGACTGGCCCTCGATACAGTAC 1800  
Qy 1801 TTCTACGTTGACTTCAGGAGTATCAATATGACCTTAAGTGGAGTTCCCGAGAGAAC 1860  
Db 1801 TTCTACGTTGACTTCAGGAGTATCAATATGACCTTAAGTGGAGTTCCCGAGAGAAC 1860  
Qy 1861 TTAGAGTTTGGGAAGTCTTGGGCTTGGGAGGTGATGAACGCCACCGCC 1920  
Db 1861 TTAGAGTTTGGGAAGTCTTGGGCTTGGGAGGTGATGAACGCCACCGCC 1920  
Qy 1921 TATGCAATTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGTGTAAAGAGAA 1980  
Db 1921 TATGCAATTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGTGTAAAGAGAA 1980  
Qy 1981 GCTGACACTGTGAAAAAGAGTCTCATCTCGAGCTCAAAATGATGACCCACCTGGGA 2040  
Db 1981 GCTGACACTGTGAAAAAGAGTCTCATCTCGAGCTCAAAATGATGACCCACCTGGGA 2040  
Qy 2041 CACCATGACACATCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCAGTGTACTTG 2100  
Db 2041 CACCATGACACATCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCAGTGTACTTG 2100  
Qy 2101 ATTTTGAATATTTGTTGCTATGCTGACCTCTCAACTACCTAAGAAAGTAAAGAGAGA 2160  
Db 2101 ATTTTGAATATTTGTTGCTATGCTGACCTCTCAACTACCTAAGAAAGTAAAGAGAGA 2160  
Qy 2161 TTTTACAGGACATGACAGAGATTTTAAAGAACATAATTTTCAAGTCTTACCTACTTTC 2220  
Db 2161 TTTTACAGGACATGACAGAGATTTTAAAGAACATAATTTTCAAGTCTTACCTACTTTC 2220  
Qy 2221 CAGGCACATTCAAATTCAGCATGCTGGTTTCCAGGAAAGTTCAGTTACACCCGCCCTTG 2280  
Db 2221 CAGGCACATTCAAATTCAGCATGCTGGTTTCCAGGAAAGTTCAGTTACACCCGCCCTTG 2280  
Qy 2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTTCTGAAGATGAGATTTGAATGAA 2340  
Db 2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTTCTGAAGATGAGATTTGAATGAA 2340  
Qy 2341 AACACAGAGAGCTGGCAGAAAGAGAGAGATTTTCAACGCTGTGAGCTTTGAGAC 2400  
Db 2341 AACACAGAGAGCTGGCAGAAAGAGAGAGATTTTCAACGCTGTGAGCTTTGAGAC 2400  
Qy 2401 CTCTCTTCTTGGCTTACCAAGTGCCTGGAATTTCTGAGTTTCAAGTCTGCT 2460  
Db 2401 CTCTCTTCTTGGCTTACCAAGTGCCTGGAATTTCTGAGTTTCAAGTCTGCT 2460  
Qy 2461 GTCCACAGAGACTGGCAGCAGGATTTGTTGGTCCACCCAGGGAAGTGTGAAGATC 2520  
Db 2461 GTCCACAGAGACTGGCAGCAGGATTTGTTGGTCCACCCAGGGAAGTGTGAAGATC 2520  
Qy 2521 TGTGACTTTGGACTGGCCCGAGACATCTGAGCGACTCCAGTACCTGTCAGGGGCAAC 2580  
Db 2521 TGTGACTTTGGACTGGCCCGAGACATCTGAGCGACTCCAGTACCTGTCAGGGGCAAC 2580

2581	Qy	GCACGGCTCCCGGTGAAGTGGATGGACCGAGAGCTTATTTGAAGGGATCTACACAATC	2540
2581	Db	GCACGGCTCCCGGTGAAGTGGATGGACCGAGAGCTTATTTGAAGGGATCTACACAATC	2540
2641	Qy	AAGAGTGACGTCGTGGTCCCTACGGCATCCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700
2641	Db	AAGAGTGACGTCGTGGTCCCTACGGCATCCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700
2701	Qy	CTTTACCTGGCANCTCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGAGTTTAAA	2760
2701	Db	CTTTACCTGGCANCTCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGAGTTTAAA	2760
2761	Qy	ATGGAGCAGCATTCATATGCCACAGAGGATATACTTTGTAATGCAATCCTCTCTGGGCT	2820
2761	Db	ATGGAGCAGCATTCATATGCCACAGAGGATATACTTTGTAATGCAATCCTCTCTGGGCT	2820
2821	Qy	TTTTGACTCAAGGAAGCGGCATCCTTCTCCCAACCTGACTTCATTTTTAGGATGTCAGCTG	2880
2821	Db	TTTTGACTCAAGGAAGCGGCATCCTTCTCCCAACCTGACTTCATTTTTAGGATGTCAGCTG	2880
2881	Qy	GCAGAGCGAAGAACGTCGTATCAGAACATCCATCCATCTACCAAAACAGCGCGCCCT	2940
2881	Db	GCAGAGCGAAGAACGTCGTATCAGAACATCCATCCATCTACCAAAACAGCGCGCCCT	2940
2941	Qy	CAGCAGAGAGCGGGCTCAGAGCCCTAGTCCCAACAGCCAGCGTGAAGATTCACAGAGAA	3000
2941	Db	CAGCAGAGAGCGGGCTCAGAGCCCTAGTCCCAACAGCCAGCGTGAAGATTCACAGAGAA	3000
3001	Qy	AGAAGTATTAGGAGAGGCGCTTGGACCCGCCACCCCTAGCAGGCTGTAGACCGCAGAGCCA	3060
3001	Db	AGAAGTATTAGGAGAGGCGCTTGGACCCGCCACCCCTAGCAGGCTGTAGACCGCAGAGCCA	3060
3061	Qy	AGATTAGCCTCGCTCTGAGAAAGCGCCCTACAGCGCTTGCCTCGCTGACATTTTCTCT	3120
3061	Db	AGATTAGCCTCGCTCTGAGAAAGCGCCCTACAGCGCTTGCCTCGCTGACATTTTCTCT	3120
3121	Qy	AGATGCTGTCTGCCATTTACTCCAAAGTGACTTCTATAAAATCAAACTCTCCTCGCACAG	3180
3121	Db	AGATGCTGTCTGCCATTTACTCCAAAGTGACTTCTATAAAATCAAACTCTCCTCGCACAG	3180
3181	Qy	GCGGGAGAGCCAATATGAGACTTGTGGTAGCCCGCCCTACCTCGGGGCGCTTCCACG	3240
3181	Db	GCGGGAGAGCCAATATGAGACTTGTGGTAGCCCGCCCTACCTCGGGGCGCTTCCACG	3240
3241	Qy	AGCTTGAGGGAAAGCCATGTATCTGAAATATAGTATATTTCTGTAATACGTGAAACAA	3300
3241	Db	AGCTTGAGGGAAAGCCATGTATCTGAAATATAGTATATTTCTGTAATACGTGAAACAA	3300
3301	Qy	ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTTTAAAATCTATGTTTTAA	3360
3301	Db	ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTTTAAAATCTATGTTTTAA	3360
3361	Qy	AATACTATGTAACTTTTTTCATCTATTAGTGATATATTTTATGGATGGAATAAATCTTTC	3420
3361	Db	AATACTATGTAACTTTTTTCATCTATTAGTGATATATTTTATGGATGGAATAAATCTTTC	3420
3421	Qy	TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA	3453
3421	Db	TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA	3453

RESULT 3					
I25169	I25169	Sequence 1 from patent US 5548065.	3453 bp	DNA	linear
LOCUS	DEFINITION				
ACCSSION	I25169				
VERSION	I25169.1	GI:1605039			
KEYWORDS					
SOURCE	unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 3453)				
	Lemischka, I.R.				

[illegible]

1981	Db	GCTGACAGCTGTGAAAAAGAGTGCTCTCATGTGCGAGGCTCAAAATGATGACCCACCCTGGGA	2040
2041	Qy	CACCATGACAAACATCGTGAATCTGCTGGGGGCATGCACACTGTGACGGCCAGTGTACTTGG	2100
2041	Db		
2041	Db	CACCATGACAAACATCGTGAATCTGCTGGGGGCATGCACACTGTGACGGCCAGTGTACTTGG	2100
2101	Qy	ATTTTGAATATGTTGCTATGTTGAGCTCCTCAACTACCTTAGAAGTAAAGAGAGAAG	2160
2101	Db		
2101	Db	ATTTTGAATATGTTGCTATGTTGAGCTCCTCAACTACCTTAGAAGTAAAGAGAGAAG	2160
2161	Qy	TTTTCAGAGGACATGGACAGAGATTTTAAAGACATAATTTCACTAGTTCCTTACCCTACTTTC	2220
2161	Db		
2161	Db	TTTTCAGAGGACATGGACAGAGATTTTAAAGACATAATTTCACTAGTTCCTTACCCTACTTTC	2220
2221	Qy	CAGGCACATTCAAATTCACGATGCCGTGTTTCACGAGAAGTTCAGTTTACACCGCCCTTG	2280
2221	Db		
2221	Db	CAGGCACATTCAAATTCACGATGCCGTGTTTCACGAGAAGTTCAGTTTACACCGCCCTTG	2280
2281	Qy	GATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATTTCTGAAGATGAGATTGAATATGAA	2340
2281	Db		
2281	Db	GATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATTTCTGAAGATGAGATTGAATATGAA	2340
2341	Qy	AACCAGAAGGCTGGCAGAGAAGAGAGGAGGAATTTGAACGTCTGACGTTTGAAGAC	2400
2341	Db		
2341	Db	AACCAGAAGGCTGGCAGAGAAGAGAGGAGGAATTTGAACGTCTGACGTTTGAAGAC	2400
2401	Qy	CTCCTTTCTTTGGGTACCAGTACCAAGTGGCCAAAGGATGAAATTCCTGGAGTTTCAAGTCGTGT	2460
2401	Db		
2401	Db	CTCCTTTCTTTGGGTACCAGTGGCCAAAGGATGGAATTCCTGGAGTTTCAAGTCGTGT	2460
2461	Qy	GTCACAGACACTGGCAGCCAGGAATGTGTGGTCAACCCACGGGAAGGTGGTCAAGATC	2520
2461	Db		
2461	Db	GTCACAGACACTGGCAGCCAGGAATGTGTGGTCAACCCACGGGAAGGTGGTCAAGATC	2520
2521	Qy	TGTCACTTTGGACTGGCCCGCAGACATCCTGAGCGCACTCCAGCTACGTCAGGGCAAC	2580
2521	Db		
2521	Db	TGTCACTTTGGACTGGCCCGCAGACATCCTGAGCGCACTCCAGCTACGTCAGGGCAAC	2580
2581	Qy	GCACGGCTGCCGTGAAGTGGATGGCACCAGAGAGCTTATTTGAAGGGATCTACACAATC	2640
2581	Db		
2581	Db	GCACGGCTGCCGTGAAGTGGATGGCACCAGAGAGCTTATTTGAAGGGATCTACACAATC	2640
2641	Qy	AAGAGTGACGTCTGGTCCCTACGGCATCCTTCTCTGGAGAGATATTTTCACTGGGTGTGAAC	2700
2641	Db		
2641	Db	AAGAGTGACGTCTGGTCCCTACGGCATCCTTCTCTGGAGAGATATTTTCACTGGGTGTGAAC	2700
2701	Qy	CCTTACCTCGCATTCCTGTGCGCACTTCTCTATAAATGATTCAGAGTGGATTTTAAA	2760
2701	Db		
2701	Db	CCTTACCTCGCATTCCTGTGCGCACTTCTCTATAAATGATTCAGAGTGGATTTTAAA	2760
2761	Qy	ATGAGCAGGCATTTCTATGCCACAGAAGGATATACTTTGTATATGCAATCCTGCTGGGCT	2820
2761	Db		
2761	Db	ATGAGCAGGCATTTCTATGCCACAGAAGGATATACTTTGTATATGCAATCCTGCTGGGCT	2820
2821	Qy	TTTGACTACAGGAAGCGGCCATCCTTCCCAACTGACTTCATTTTTAGGATGTCAGCTG	2880
2821	Db		
2821	Db	TTTGACTACAGGAAGCGGCCATCCTTCCCAACTGACTTCATTTTTAGGATGTCAGCTG	2880
2881	Qy	GCAGAGCAGAGAGGATGTATCAGAAACATCCATCCATCTACCAAAACAGGCGGCCCT	2940
2881	Db		
2881	Db	GCAGAGCAGAGAGGATGTATCAGAAACATCCATCCATCTACCAAAACAGGCGGCCCT	2940
2941	Qy	CAGCAGAGAGCGGGCTCAGAGCCAGTCGCCACAGCGCCAGGTGAAGATTCACAGAGAA	3000
2941	Db		
2941	Db	CAGCAGAGAGCGGGCTCAGAGCCAGTCGCCACAGCGCCAGGTGAAGATTCACAGAGAA	3000
3001	Qy	AGAAAGTTAGCGAGGAGCGCTTGGACCCCGCCACCCCTAGCAGGCTGTAGACCGCAGAGCCA	3060
3001	Db		
3001	Db	AGAAAGTTAGCGAGGAGCGCTTGGACCCCGCCACCCCTAGCAGGCTGTAGACCGCAGAGCCA	3060
3061	Qy	AGATTAGCTCGCCTCTGAGGAAGCGCCCTACAGCGGCTTGTCTTCGCTGCACTTTTCTCT	3120
3061	Db		
3061	Db	AGATTAGCTCGCCTCTGAGGAAGCGCCCTACAGCGGCTTGTCTTCGCTGCACTTTTCTCT	3120





1441 AAATCTCCAAATTCACGAGGAAATCCAGAAAGAGTTGGAAATAAAAAAGCTAACAGA 1500  
1441 AAATCTCCAAATTCACGAGGAAATCCAGAAAGAGTTGGAAATAAAAAAGCTAACAGA 1500  
1501 AAAGTGTGGCAGTGGGTGTCAGAGTACTCTAAATATAGTGAAGCCGGGAAGGG 1560  
1501 AAAGTGTGGCAGTGGGTGTCAGAGTACTCTAAATATAGTGAAGCCGGGAAGGG 1560  
1561 CTCTGTCGTCATGTCGCTGACAAATCTATGGCAGCTCTTGGAAACCATCTTTTA 1620  
1561 CTCTGTCGTCATGTCGCTGACAAATCTATGGCAGCTCTTGGAAACCATCTTTTA 1620  
1621 AACTCACCAGGCCCTTCCCTTTTCATCCAAAGAACATCTCTTATGCGACCATGGG 1680  
1621 AACTCACCAGGCCCTTCCCTTTTCATCCAAAGAACATCTCTTATGCGACCATGGG 1680  
1681 CTCTGTCGTCATGTCGCTGACAAATCTATGGCAGCTCTTGGAAACCATCTTTTA 1740  
1681 CTCTGTCGTCATGTCGCTGACAAATCTATGGCAGCTCTTGGAAACCATCTTTTA 1740  
1741 TTTAGGTACGAGTACGCTGACGATGATCCAGGTGACTGGCCCTGGATACAGGTAC 1800  
1741 TTTAGGTACGAGTACGCTGACGATGATCCAGGTGACTGGCCCTGGATACAGGTAC 1800  
1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAAGTGGAGTTCCTCCGAGAGAAC 1860  
1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAAGTGGAGTTCCTCCGAGAGAAC 1860  
1861 TTAGAGTTGGAGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTC 1920  
1861 TTAGAGTTGGAGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTCCTGGGTC 1920  
1921 TATGTCATGTAAGAACGGAGTCTCAATTCAGGTGGCGTCAAGATGCTTAAAGAGAA 1980  
1921 TATGTCATGTAAGAACGGAGTCTCAATTCAGGTGGCGTCAAGATGCTTAAAGAGAA 1980  
1981 GCTGACAGCTGTGAAAAAGAGTCTCATGTCGAGCTCAAAATGATGACCCACCTGGGA 2040  
1981 GCTGACAGCTGTGAAAAAGAGTCTCATGTCGAGCTCAAAATGATGACCCACCTGGGA 2040  
2041 CACATGACACATCGTAATCTGCTGGGGGATGCACACTGTAGGCGCAGTGTACTTG 2100  
2041 CACATGACACATCGTAATCTGCTGGGGGATGCACACTGTAGGCGCAGTGTACTTG 2100  
2101 ATTTTGAATATGTTGTCATGTCGCTGACCTCCTCAACTACCTAAGAGTAAAGAGAGAG 2160  
2101 ATTTTGAATATGTTGTCATGTCGCTGACCTCCTCAACTACCTAAGAGTAAAGAGAGAG 2160  
2161 TTTACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCAGTTCCTTACCTACTTTC 2220  
2161 TTTACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCAGTTCCTTACCTACTTTC 2220  
2221 CAGGACATTCAAATTCAGCATGCTGGTTCACGAGAGTTTCAGTTACACCCGCCCTTG 2280  
2221 CAGGACATTCAAATTCAGCATGCTGGTTCACGAGAGTTTCAGTTACACCCGCCCTTG 2280  
2281 GATCAGCTCAGGCTTCAATGGGAATTCATTCATCTGAGATGAGATTCGAATATGAA 2340  
2281 GATCAGCTCAGGCTTCAATGGGAATTCATTCATCTGAGATGAGATTCGAATATGAA 2340  
2341 AACACAGAGAGGCTGGCAGAGAGAGAGAGAGATTTGAACGTCGCTGACGTTTGAAGAC 2400  
2341 AACACAGAGAGGCTGGCAGAGAGAGAGAGAGATTTGAACGTCGCTGACGTTTGAAGAC 2400  
2401 CTCTCTTGTGTCGTCACGAGTGGCCAAAGCATGGAATTCCTGGAGTTCAGTCTGCT 2460  
2461 GTCCACAGAGACCTGGCAGCAGGAATGTGTTGTCACCCAGGAAAGGTGTGAAGATC 2520  
2461 GTCCACAGAGACCTGGCAGCAGGAATGTGTTGTCACCCAGGAAAGGTGTGAAGATC 2520  
2521 TGTGACTTGGACTGGCCCGAGACATCTCTGACGACTCCAGTACGTCGTCAGGGGCAC 2580

2521 TGTGACTTGGACTGGCCCGAGACATCTCTGACGACTCCAGCTACGTCGTCAGGGGCAC 2580  
2581 GCACGGCTGCCGTCGAGTGGATGGACCCGAGAGCTTATTTGAAGGATCTACACAATC 2640  
2581 GCACGGCTGCCGTCGAGTGGATGGACCCGAGAGCTTATTTGAAGGATCTACACAATC 2640  
2641 AAGAGTGAGCTCTGGTCCTACGSCATCTCTCTGGGAGATATTTTACATGGGTGTGAAC 2700  
2641 AAGAGTGAGCTCTGGTCCTACGSCATCTCTCTGGGAGATATTTTACATGGGTGTGAAC 2700  
2701 CTTTACCTTGGCATCTCTCTGACGCTAACTTCTATAAACTGATTTCAGAGTGAATTTAA 2760  
2701 CTTTACCTTGGCATCTCTCTGACGCTAACTTCTATAAACTGATTTCAGAGTGAATTTAA 2760  
2761 ATGGAGCAGCATCTATGCCACAGAGGATATCTTTGTATGTAATCTCTCTGGCT 2820  
2761 ATGGAGCAGCATCTATGCCACAGAGGATATCTTTGTATGTAATCTCTCTGGCT 2820  
2821 TTTGACTCAAGGAAGCGGCATCTTCCCACTGACTTCATTTTAGGATGTCAGCTG 2880  
2821 TTTGACTCAAGGAAGCGGCATCTTCCCACTGACTTCATTTTAGGATGTCAGCTG 2880  
2881 GCAGAGGAGAGAGAGCATGTATCAGAACATCCATCCATACCAAAACAGCGGCCCT 2940  
2881 GCAGAGGAGAGAGAGCATGTATCAGAACATCCATCCATACCAAAACAGCGGCCCT 2940  
2941 CAGCAGAGAGGCGGCTCAGAGCCAGTCCACACAGCGGCGAGTGAAGATTCACAGAA 3000  
2941 CAGCAGAGAGGCGGCTCAGAGCCAGTCCACACAGCGGCGAGTGAAGATTCACAGAA 3000  
3001 AGAAGTTACGAGAGGCGCTTGGACCCGCGCTAGCAGGCTGTAGACCGCAGAGCCA 3060  
3001 AGAAGTTACGAGAGGCGCTTGGACCCGCGCTAGCAGGCTGTAGACCGCAGAGCCA 3060  
3061 AGATTAGCCTCGCTCTGAGGAAGCGCTACAGCGCTGTCTGCTGGACTTTTCTCT 3120  
3061 AGATTAGCCTCGCTCTGAGGAAGCGCTACAGCGCTGTCTGCTGGACTTTTCTCT 3120  
3121 AGATGCTGTCTGCCATTACTCCAAAGTACTTCTATAAAATCAAACTCTCTCGCACAG 3180  
3121 AGATGCTGTCTGCCATTACTCCAAAGTACTTCTATAAAATCAAACTCTCTCGCACAG 3180  
3181 GCGGAGAGCCAAATATGAGACTTGTGTGAGCCGCGCTTACCTGGGGCGCTTTCACG 3240  
3181 GCGGAGAGCCAAATATGAGACTTGTGTGAGCCGCGCTTACCTGGGGCGCTTTCACG 3240  
3241 AGCTTGAGGGGAAAGCCATGTATCTGAAATATAGTATATTCCTGTAATACGTGAACAA 3300  
3241 AGCTTGAGGGGAAAGCCATGTATCTGAAATATAGTATATTCCTGTAATACGTGAACAA 3300  
3301 ACCAAACCCGTTTTCGCTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTAA 3360  
3301 ACCAAACCCGTTTTCGCTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTAA 3360  
3361 AATACATATCTAATTTTTCATCTATTTAGTGATATATTTATGATGGAATAAATTC 3420  
3361 AATACATATCTAATTTTTCATCTATTTAGTGATATATTTATGATGGAATAAATTC 3420  
3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAA 3453  
3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAA 3453

## RESULT 5

MUSFLK2  
LOCUS MUSFLK2 3426 bp mRNA linear ROD 27-APR-1993  
DEFINITION Mouse flk-2 mRNA, complete cds.  
ACCESSION M64689  
VERSION M64689.1 GI:193327  
KEYWORDS Mus musculus (strain C3H/He) cDNA to mRNA.  
SOURCE Mus musculus









Db 2032 GCTGACAGCTGTGAAAGAAAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2091  
Qy 2041 CACCATGACAACTCGTGAATCTGCTGGGGCATGACACACTGTGAGGCCAGTGTACTTG 2100  
Db 2092 CACCATGACAACTCGTGAATCTGCTGGGGCATGACACACTGTGAGGCCAGTGTACTTG 2151  
Qy 2101 ATTTTGAATATTTGCTGATGGTGAACCTCTCACTACCTAAAGAAAGTAAAGAGAGAG 2160  
Db 2152 ATTTTGAATATTTGCTGATGGTGAACCTCTCACTACCTAAAGAAAGTAAAGAGAGAG 2211  
Qy 2161 TTTCACAGGACATGACAGAGATTTTAAAGAAACATATTTTCACTTCACTTCTTCC 2220  
Db 2212 TTTCACAGGACATGACAGAGATTTTAAAGAAACATATTTTCACTTCACTTCTTCC 2271  
Qy 2221 CAGCACAATTCAAATTCACAGATGCTGTTTCAAGAGAGTTCAGTTACACCCGCCCTTG 2280  
Db 2272 CAGCACAATTCAAATTCACAGATGCTGTTTCAAGAGAGTTCAGTTACACCCGCCCTTG 2331  
Qy 2281 GATCAGCTCTCAGGGTTCAATGGGAATTCATCTGAAGATGAGATTGAATATGAA 2340  
Db 2332 GATCAGCTCTCAGGGTTCAATGGGAATTCATCTGAAGATGAGATTGAATATGAA 2391  
Qy 2341 AACCAGAGAGGCTGGCAGAGAGAGAGAGAGATTTGAACGCTCTGACGTTTGAAGAC 2400  
Db 2392 AACCAGAGAGGCTGGCAGAGAGAGAGAGAGATTTGAACGCTCTGACGTTTGAAGAC 2451  
Qy 2401 CTCCTTTGCTTTGGCTACCAAGTGCCAAAGSCATGGAATTCCTGGAGTTCAGTCGTGT 2460  
Db 2452 CTCCTTTGCTTTGGCTACCAAGTGCCAAAGSCATGGAATTCCTGGAGTTCAGTCGTGT 2511  
Qy 2461 GTCCACAGAGACCTGGCAGCAGGAATGTGTTGTCACCCAGCGGAAGTGGTGAAGATC 2520  
Db 2512 GTCCACAGAGACCTGGCAGCAGGAATGTGTTGTCACCCAGCGGAAGTGGTGAAGATC 2571  
Qy 2521 TGTGACTTTGACTGGCCCGGAGACATCTGAGCGACTCCAGTCTGTCAGGGGCAAC 2580  
Db 2572 TGTGACTTTGACTGGCCCGGAGACATCTGAGCGACTCCAGTCTGTCAGGGGCAAC 2631  
Qy 2581 GCAGGCTGCGGTGAAGTGGATGSCACCCAGAGCTTATTTGAAGGATCTACACAATC 2640  
Db 2632 GCAGGCTGCGGTGAAGTGGATGSCACCCAGAGCTTATTTGAAGGATCTACACAATC 2691  
Qy 2641 AAGAGTACGCTCTGCTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700  
Db 2692 AAGAGTACGCTCTGCTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2751  
Qy 2701 CCTTACCTGGCATCTCTGTCAGCCTAACTTCTATAAAGTATTCAGAGTGGATTTAA 2760  
Db 2752 CCTTACCTGGCATCTCTGTCAGCCTAACTTCTATAAAGTATTCAGAGTGGATTTAA 2811  
Qy 2761 ATGAGCAGCATTCTATGCCACAGAGGATATATTTGTAATCAATCCTGCTGGGCT 2820  
Db 2812 ATGAGCAGCATTCTATGCCACAGAGGATATATTTGTAATCAATCCTGCTGGGCT 2871  
Qy 2821 TTGACTCAAGGAAGCGCCATCTTCCCAACCTGACTTCTATTTTAGGATGTCAGCTG 2880  
Db 2872 TTGACTCAAGGAAGCGCCATCTTCCCAACCTGACTTCTATTTTAGGATGTCAGCTG 2931  
Qy 2881 GCAGAGCAGAGAGAGC-----ATGATACAGAACATCCATCC 2917  
Db 2932 GCAGAGCAGAGAGAGAGCGATGATCAGAACATGGGTGGCAAGCTCCCAAGAACATCCATCC 2991  
Qy 2918 ATCTACCAAAACAGCGCGCCCTCAGCAGAGAGCGGCTCAGAGCC-CAGTCGCCACAG 2976  
Db 2992 ATCTACCAAAACAGCGCGCCCTCAGCAGAGAGCGGCTCAGAGCCCGCATCCGACAG 3051  
Qy 2977 CGCCAGGTGAAGATTCACAGAGAAAGAGTGTAGCGAGGAGCGCTTGGACCCGCCACCT 3036  
Db 3052 GCCCAGGTGAAGATTCACAGAGAAAGAGTGTAGCGAGAGCGCTTGGACCCGCCACCT 3111  
Qy 3037 AGCAGGCTGTAGACCGCAGAGCAAGATTAGCCTCTGAGAGAGCGCCCTACAGCG 3096  
Db 3112 AGCAGGCTGTAGACCGCAGAGCAAGATTAGCCTCTGAGAGAGCGCCCTACAGCG 3171

Qy 3097 CGTGTCTCGCTGGAGCTTTTCTCTAGATGCTGTCTGCCATTACTCCAAAGTGACTTCTAT 3156  
Db 3172 CGTGTCTCGCTGGAGCTTTTCTCTAGATGCTGTCTGCCATTACTCCAAAGTGACTTCTAT 3231  
Qy 3157 AAAATCAAACTCTCTCTGCACAGCGGGAGAGCCAAATATGAGACTTGTGTGTAGCCCC 3216  
Db 3232 AAAATCAAACTCTCTCTGCACAGCGGGAGAGCCAAATATGAGACTTGTGTGTAGCCCC 3291  
Qy 3217 GCCTACCTGGGGGCTTT-----CCACGAGCTTGAGGGAAAGCCATGATCTGAA 3268  
Db 3292 GCCTACCTGGGGGCTTTCCAGGCCCTCCAGGCTTGAGGGAAAGCCATGATCTGAA 3351  
Qy 3269 ATATAGTATATTTCTGTAATACGTGAACAAACCAACCCCGTTTTTGTAAAGGAAAG 3328  
Db 3352 ATATAGTATATTTCTGTAATACGTGAACAAACCAACCCCGTTTTTGTAAAGGAAAG 3411  
Qy 3329 CTAAATATGATTTTAAATATCTATGTTTAAATCTATGTAACCTTTTTCATCTATTTA 3388  
Db 3412 CTAAATATGATTTTAAATATCTATGTTTAAATCTATGTAACCTTTTTCATCTATTTA 3471  
Qy 3389 GTGATATATTTTATGATGGAATAAATCTTCTACTGTAAATAAAAAA 3438  
Db 3472 GTGATATATTTTATGATGGAATAAATCTTCTACTGTAAATAAAAAA 3521

## RESULT 7

MMFLT3  
LOCUS MMFLT3 3521 bp mRNA linear ROD 04-DEC-1991  
DEFINITION Mouse Flt3 mRNA for tyrosine kinase receptor of the PDGF.  
ACCESSION X59398  
VERSION X59398.1 GI:50978  
KEYWORDS Flt3 gene; tyrosine kinase receptor.  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
REFERENCE 1 (bases 1 to 1321)  
AUTHORS Rosnet O.  
TITLE Direct Submission  
JOURNAL Submitted (07-MAY-1991) O. Rosnet, Cancerologie et therapeutique exp., Inst National de la Sante, 27 Boulevard Lei Roure, 13009 Marseille, France  
REFERENCE 2 (bases 1 to 3521)  
AUTHORS Rosnet O., Marchetto, S., delapeyriere, O. and Birnbaum, D.  
TITLE Murine Flt3, a gene encoding a novel tyrosine kinase receptor of the PDGFR/CSF1R family  
JOURNAL Oncogene 6 (9), 1641-1650 (1991)  
MEDLINE 92019834  
PUBMED 1656368

## FEATURES

Location/Qualifiers  
1..3521  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
1..3521  
/gene="Flt3"  
1..3521  
/gene="Flt3"  
/evidence=experimental  
82..3084  
/gene="Flt3"  
/codon\_start=1  
/product="Flt3 protein"  
/protein\_id="CAA42041.1"  
/db\_xref="GI:50979"  
/db\_xref="MGI:95559"  
/translation="MRAAQRSDRLLLVLSVILETVTNODLPVVKCVLISHENN  
GSSACKPSYRMVGRSPEDLOCTPRQSEGTVEATVEAESGSLTLOVALTPDGL  
SCIWFKHSSLCQPHFDIQNRGIVSMALINVTETQAGETLLHQISEANVTYLVFN  
VDTQLYLVRRPYFRKNMDALLCISIEGVPETVNEWLCSSHRECKEKGPAVRKE  
EKLVELFGTDIRCCARNALGRESTKLTIDLNQAPQSLQFLKVGEPFLTRKAI  
HVNHGFLTWELEDKALEBGSYFENSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSS

943 a 854 c 907 g 817 t

Query Match	96.8%;	Score 3344;	DB 10;	Length 3521;
Best Local Similarity	98.5%;	Pred. No. 0;		
Matches 3418;	Conservative	0;	Mismatches 20;	Indels 32; Gaps 3;
Qy	1	CGGCGCTGCTACCGCGCTCCGGAGCCATCGGGCGTTGGCGCAGCCAGCGACCGG	60	
Db	52	CGGCGCTGCTACCGCGCTCCGGAGCCATCGGGCGTTGGCGCAGCCAGCGACCGG	111	
Qy	61	CGGCTGCTGCTGTTGTTTGTCTAGTAATGATTTTGAGACCGTTACAAACCAAGAC	120	
Db	112	CGGCTGCTGCTGTTGTTTGTCTAGTAATGATTTTGAGACCGTTACAAACCAAGAC	171	
Qy	121	CTGCTGTGATCAAGTGTGTTTAACTCAGTCATGAGAACAAATGCTCATCAGCGGGAAG	180	
Db	172	CTGCTGTGATCAAGTGTGTTTAACTCAGTCATGAGAACAAATGCTCATCAGCGGGAAG	231	
Qy	181	CCATCATCGTACCGAATGTCGAGGATCCCAGAACCTCCAGTCATCCCGAGGCGC	240	
Db	232	CCATCATCGTACCGAATGTCGAGGATCCCAGAACCTCCAGTCATCCCGAGGCGC	291	
Qy	241	CAGAGTGAAGGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCGCGAGTCTGGTCCATC	300	
Db	292	CAGAGTGAAGGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCGCGAGTCTGGTCCATC	351	
Qy	301	ACCTCGCAAGTGCAGCTGCCACCCAGGGAGCTTTCTGCTGCTCGGTCTTTAAGCAC	360	
Db	352	ACCTCGCAAGTGCAGCTGCCACCCAGGGAGCTTTCTGCTGCTCGGTCTTTAAGCAC	411	
Qy	361	AGCTCCCTGGGCTGCCAGCGCACTTTGATTTACAAAACAGAGAAATCGTTTCCATGCC	420	
Db	412	AGCTCCCTGGGCTGCCAGCGCACTTTGATTTACAAAACAGAGAAATCGTTTCCATGCC	471	
Qy	421	ATCTTGAACTGCAGAGACCCAGGCAAGAGAAATACCTACTCCATATTCAGAGCGAAGC	480	
Db	472	ATCTTGAACTGCAGAGACCCAGGCAAGAGAAATACCTACTCCATATTCAGAGCGAAGC	531	
Qy	481	GCCAACTACACAGTACTGTTTCACAGTCAATGTATAGATACACAGCTGTATGCTAAGG	540	
Db	532	GCCAACTACACAGTACTGTTTCACAGTCAATGTATAGATACACAGCTGTATGCTAAGG	591	
Qy	541	AGACCTTTACTTTAGGAAGATGAAAACAGGATGCATGCTCTGCATCTCCGAGGGTGT	600	
Db	592	AGACCTTTACTTTAGGAAGATGAAAACAGGATGCATGCTCTGCATCTCCGAGGGTGT	651	
Qy	601	CCGAGAGCCCTGTGGAGTGGTGCTCTGCAGCTCCACAGGGGAAAGCTGTAAAGAGAA	660	
Db	652	CCGAGAGCCCTGTGGAGTGGTGCTCTGCAGCTCCACAGGGGAAAGCTGTAAAGAGAA	711	
Qy	661	GGCCCTGCTGTTTCAGAAAGGAGGAAAAGGTACTTTCATGATGTTTCCGACACAGATC	720	
Db	712	GGCCCTGCTGTTTCAGAAAGGAGGAAAAGGTACTTTCATGATGTTTCCGACACAGATC	771	
Qy	721	AGATGCTGTGCTAGAAATGCACCTGGGCGGGAATGCACCAAGCTGTTCCACATAGTCTA	780	
Db	772	AGATGCTGTGCTAGAAATGCACCTGGGCGGGAATGCACCAAGCTGTTCCACATAGTCTA	831	
Qy	781	AACCAAGCTCCTCAGAGCACACTGCCCCAGTTATTTCTGAAAGTGGGGGAACCCCTTGTG	840	

QY 1921 TATGGCATTAGTAAACGGGAGTCTCAATTCAGTGGCGTGAAGATGCTAAAAAGAGAA 1980  
DB 1972 TATGGCATTAGTAAACGGGAGTCTCAATTCAGTGGCGTGAAGATGCTAAAAAGAGAA 2031  
QY 1981 GCTGACAGCTGTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGAACCCAGCTGGGA 2040  
DB 2032 GCTGACAGCTGTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGAACCCAGCTGGGA 2091  
QY 2041 CACCATGACACATCGTGAATCTGCTGGGCGATGCACACTGTCAGGCGCAGTGTACTTG 2100  
DB 2092 CACCATGACACATCGTGAATCTGCTGGGCGATGCACACTGTCAGGCGCAGTGTACTTG 2151  
QY 2101 ATTTTCAATATTCTGCTATGTTGAGCTCTCCTCAACTACCTAAGAAGTAAAGAGAGAA 2160  
DB 2152 ATTTTCAATATTCTGCTATGTTGAGCTCTCCTCAACTACCTAAGAAGTAAAGAGAGAA 2211  
QY 2161 TTTCACAGGACATGGACAGAGATTTTAAAGAAACATAATTTACCTTACCTTACTTTC 2220  
DB 2212 TTTCACAGGACATGGACAGAGATTTTAAAGAAACATAATTTACCTTACCTTACTTTC 2271  
QY 2221 CAGGCACATTCAATTTCCAGATCCCTGGTTCAGGAGAGTTCAGTTACACCCGCCCTTG 2280  
DB 2272 CAGGCACATTCAATTTCCAGATCCCTGGTTCAGGAGAGTTCAGTTACACCCGCCCTTG 2331  
QY 2281 GATCAGCTCTCAGGTTTCAATGGGAATTCATTTCTGAAGATGAGATTTGAATATGAA 2340  
DB 2332 GATCAGCTCTCAGGTTTCAATGGGAATTCATTTCTGAAGATGAGATTTGAATATGAA 2391  
QY 2341 AACCAGAGAGGCTGGCAGAGAGAGAGAGAGAGATTTGAACGTCGTCAGCTTTGAAGAC 2400  
DB 2392 AACCAGAGAGGCTGGCAGAGAGAGAGAGAGAGATTTGAACGTCGTCAGCTTTGAAGAC 2451  
QY 2401 CTCCTTTGCTTTGCTTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCGNGT 2460  
DB 2452 CTCCTTTGCTTTGCTTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCGNGT 2511  
QY 2461 GTCCACAGAGACCTGGCAGCAGGAATGTGTGTGTCACCCAGGAGAGTGGTGAAGATC 2520  
DB 2512 GTCCACAGAGACCTGGCAGCAGGAATGTGTGTGTCACCCAGGAGAGTGGTGAAGATC 2571  
QY 2521 TGTGACTTTGAGCTGGCCCGAGAGATCCTGAGCGACTCCAGCTACGTCGTCAGGGGCAAC 2580  
DB 2572 TGTGACTTTGAGCTGGCCCGAGAGATCCTGAGCGACTCCAGCTACGTCGTCAGGGGCAAC 2631  
QY 2581 GCAGGCTGCGGTGAAGTGAATGGCAGCAGAGCTTATTTGAGGAGATCTACACAATC 2640  
DB 2632 GCAGGCTGCGGTGAAGTGAATGGCAGCAGAGCTTATTTGAGGAGATCTACACAATC 2691  
QY 2641 AAGAGTACGCTGTGCTCTACGGCATCCTCTCGGGAGATATTTTCACTGGGTGTGAAC 2700  
DB 2692 AAGAGTACGCTGTGCTCTACGGCATCCTCTCGGGAGATATTTTCACTGGGTGTGAAC 2751  
QY 2701 CCTTACCCTGGCATTCTGTCGAGCTCAACTTCTATAAAGTATTCAGAGTGAATTTAAA 2760  
DB 2752 CCTTACCCTGGCATTCTGTCGAGCTCAACTTCTATAAAGTATTCAGAGTGAATTTAAA 2811  
QY 2761 ATGAGAGAGCATTCTATGTCACAGAGAGGATATCTTTGTAATCCATCTGCTGGGCT 2820  
DB 2812 ATGAGAGAGCATTCTATGTCACAGAGAGGATATCTTTGTAATCCATCTGCTGGGCT 2871  
QY 2821 TTTGACTCAAGAGAGCGGCATCTCTCCCAAGCTGACTTCTATTTTGGAGTGTGAGCTG 2880  
DB 2872 TTTGACTCAAGAGAGCGGCATCTCTCCCAAGCTGACTTCTATTTTGGAGTGTGAGCTG 2931  
QY 2881 GCAGAGGAGAGAGAGC-----ATGTCATCAGAAATCCATCC 2917  
DB 2932 GCAGAGGAGAGAGAGAGATGTATCAGAAATGGGTGGCAACGTCCTCCAGAAATCCATCC 2991  
QY 2918 ATCTACCAAAACAGCGGCCCTCTCAGCAGAGAGCGGGCTCAGAGC-CAGTGGCCACAG 2976  
DB 2992 ATCTACCAAAACAGCGGCCCTCTCAGCAGAGAGCGGGCTCAGAGCGGCCCTCAGCCACAG 3051

QY 2977 GCACAGGTGAAGATTCACAGAGAAAGATTAGCAGGAGGCTTTGGACCCCGCCACCCT 3036  
DB 3052 GCCCAGGTGAAGATTCACAGAGAAAGATTAGCAGGAGGCTTTGGACCCCGCCACCCT 3111  
QY 3037 AGCAGGCTGTAGACCGCAGAGCCAGATTAGCCTCGCCTCTCAGGAAGCGCCTTACAGCG 3096  
DB 3112 AGCAGGCTGTAGACCGCAGAGCCAGATTAGCCTCGCCTCTCAGGAAGCGCCTTACAGCG 3171  
QY 3097 CGTTGCTTCGCTGGACTTTTCTAGATGCTGTCTGCCATTACTCCAAAGTGAATCTTAT 3156  
DB 3172 CGTTGCTTCGCTGGACTTTTCTAGATGCTGTCTGCCATTACTCCAAAGTGAATCTTAT 3231  
QY 3157 AAAATCAAACTCTCCTCGCAGCGGAGAGCCAAATAATAGACTTGTGGTGAAGCC 3216  
DB 3232 AAAATCAAACTCTCCTCGCAGCGGAGAGCCAAATAATAGACTTGTGGTGAAGCC 3291  
QY 3217 GCCTACCTCGGGGCTTT-----CCACGAGCTTGAGGGAAGCCATGATCTGAA 3268  
DB 3292 GCCTACCTCGGGGCTTTTCCAGGCCCTTCCAGGCTTGAGGGAAGCCATGATCTGAA 3351  
QY 3269 ATATAGTATATTTTGTAAATACGTGAACCAACCAACCCCTTTTCTGAAGGAAAG 3328  
DB 3352 ATATAGTATATTTTGTAAATACGTGAACCAACCAACCCCTTTTCTGAAGGAAAG 3411  
QY 3329 CTAATATGATTTTAAATACTATGTTTAAATACTATGTAACCTTTTCTATCTATTTA 3388  
DB 3412 CTAATATGATTTTAAATACTATGTTTAAATACTATGTAACCTTTTCTATCTATTTA 3471  
QY 3389 GTGATATATTTATGATGGAATAAATCTTCTACTGTAAAAAATAAATAA 3438  
DB 3472 GTGATATATTTATGATGGAATAAATCTTCTACTGTAAAAAATAAATAA 3521

RESULT 8  
AR005212  
LOCUS AR005212 3501 bp DNA linear PAT 04-DEC-1998  
DEFINITION Sequence 3 from patent US 5747651.  
ACCESSION AR005212  
VERSION AR005212.1 GI:3966091  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3501)  
AUTHORS Lemischka, I.R.  
TITLE Antibodies against tyrosine kinase receptor flk-1  
JOURNAL Patent: US 5747651-A 3 05-MAY-1998;  
FEATURES  
source location/Qualifiers  
I. 3501  
BASE COUNT 1068 a 709 c 784 g 940 t  
ORIGIN

Query Match 64.5%; Score 2227.2; DB 6; Length 3501;  
Best Local Similarity 79.6%; Pred. No. 0;  
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;  
QY 8 GCCTACCGCGCTCGGAGGCCATCGGGCGTTGGCGAGCGCAGCGCGGCGTGC 67  
DB 35 GGGGACCCCGGCTCGGAGGCCATCGCGCGTTGGC---GCAGCGCGGCGCACCGTGC 91  
QY 68 TGCTGCTGCTGTTTGTTCAGTAATGATTTCTTGAGACCGTTACAAACCAAGACTGCCTG 127  
DB 92 CGCTGCTGCTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCTG 151  
QY 128 TGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGGAAGCCATCAT 187  
DB 152 TGATCAAGTGTGTTTAAATCAATCAATAAGAACAAATGATTCATCAGTGGGAAAGTCATCAT 211  
QY 188 CGTACCGAATGTTGCGGAGGATCCCCAGAGACCTCCAGTGTACCCCGAGCGGACAGAGTG 247  
DB 212 CATATCCCATGTTATCAGAAATCCCGGAAGACCTCGGGTGTGCTTGAGACCCAGAGCT 271



QY 248 AAGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATCACCCCTGC 307  
Db 272 CAGGGACAGTGTACGAAGCTGCCGCTGTGGAAGTGGATGTATCTGCTTCCATCACACTGC 331  
QY 308 AAGTCAGCTCGCCACCCAGGAGGACCTTTCCTGCTCTGGGTCTTTAAGCACAGCTGCC 367  
Db 332 AAGTCGTGTGATGTCGCCAGGGAACATTTCTGTCTCTGGGTCTTTAAGCACAGCTGCC 391  
QY 368 TGGGCTGCCAGCGCACTTTGATTTTACAAACAGAGAAATCGTTCATCGCCATCTTGA 427  
Db 392 TGAATTTGCCAGCCACATTTTGAATTTACAAACAGAGAGTGTGTTTCCATGGTCAATTTGA 451  
QY 428 ACCTGACAGACCCAGGAGGAGATACCTACTCCATATTCAGAGCGAAGCGCCAACT 487  
Db 452 AATGACAGAAACCCAGCTGGAGATACCTACTTTTATTCAGAGTGAAGCTACCAAT 511  
QY 488 ACACAGTACTGTTACAGTGGAATGTAAAGAGATACACAGCTGTATGTCTAAGGAGACCT 547  
Db 512 ACACAATATTTTACAGTGAGTAAAGAAATACCCCTGCTTTACACATTAAGAAGACCTT 571  
QY 548 ACTTTAGGAAGATGGAACACAGATGCACTGCTCTGCATCTCCGAGGCTGTTCCGAGC 607  
Db 572 ACTTTAGAAAATGGAACACAGGACCGCCCTGCTGCATATCTGAGAGCTTCCAGAGC 631  
QY 608 CCACTGTGGAGTGGTCTCTGAGCTCCACAGGGAAGCTGTAAAGAAAGAGGCGCCTG 667  
Db 632 CGATCGTGGATGGTCTTTGGATTCACAGGGGGAAGCTGTAAAGAAAGAGTCCAG 691  
QY 668 CTGTTGTCAGAAAGGAGGAAAGTACTTTCATGATGTTTCCGGAACAGACATCAGATGCT 727  
Db 692 CTGTTGTTAAAGAGGAGGAAAGTGTCTTCATGAATTTATTTGGACGGACATAGGTGCT 751  
QY 728 GTGCTAGAAATGCACTGGGCGCGAATGCACAGCTGTTTCCACATAGATCTAAACAGG 787  
Db 752 GTGCCAGAAATGAATGGGAGGGAATGCACAGGCTGTTTCCACATAGATCTAAATCAA 811  
QY 788 CTCTCAGAGCACACTGCCCCAGTATTCTGAAAGTGGGGAACCTTGTGGATCAGT 847  
Db 812 CTCTCAGACCACATTTGCCACAATTTCTTAAAGTAGGGGAACCTTTATGGTAAGGT 871  
QY 848 GTAAGGCCATCCATGTGAACCATGGATTCGGGCTCACTTGGGAGCTGGAAGACAAAGCCC 907  
Db 872 GCAAGCTGTTCATGTGAACCATGGATTCGGGCTCACTTGGGAATTAGAAACAAAGCAC 931  
QY 908 TGAGAGGCGCAGCTACTTTGAGATGAGTACCTACTCCACAACAGACCATGATTCGGA 967  
Db 932 TCAGAGGCGCACTACTTTGAGATGAGTACCTATTCAACAAACAGAACTATGATACGGA 991  
QY 968 TTCTCTTGGCCTTTGTCTCTCGTGGGAAGGAACGACACCGGATATTACACCTGCTCTT 1027  
Db 992 TTCTGTTTGTGTTATCATCATGAGTGGAAGAAACGACACCGGATATACACTTGTCTCT 1051  
QY 1028 CCTCAAGACCCAGCCAGCTAGCTGTGGTGGACCATCTTAGAAAAAGGGTTTATAACG 1087  
Db 1052 CTTCAAGACATCCAGTCAATCAGCTTTGGTTACCATCGTAGGAAGGGATTTATAATG 1111  
QY 1088 CTACAGCTCGAAGAAAGATGAAATTTGACCGGTAGCAAAAGTCTTCTCTTCAGTCA 1147  
Db 1112 CTACCAATTTCAAGTGAAGTATGAAATTTGACCAATATGAAGAGTTTTGTCTCTGCA 1171  
QY 1148 GGTTTAAAGCTACCCAGCAATCCGATGCACTGAGTCTCTCTCAAGCCTCATTTCCCT 1207  
Db 1172 GGTTTAAAGCTACCCACAATCAGATGTAGTGGACCTTCTCTCGAANAATCATTTCCCT 1231  
QY 1208 GTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCATTAAGAACA 1267  
Db 1232 GTGAGCAAAAGGCTTTGATACGGATACAGCATATCCAGTTTGGCAATCATAAAGCAC 1291  
QY 1268 AGCCAGGAGTACATATTTCTGACAGAAATGATGACGCCAGTTCACCAAAATGTCA 1327  
Db 1292 AGCCAGGAGATATATATTCATGTCAGCAAAATGATGATGCCAATTTACCAAAATGTCA 1351  
QY 1328 CGCTGAATATGAAGAAACCTCAAGTGTAGCAAAATGCTTCAGCCAGCCAGCGCTCCT 1387

Db 1352 CGCTGAATATGAAGAAACCTCAAGTGTCTCGGAAGAGCATCGGCAAGTTCAGGCGTCT 1411  
QY 1388 GTTCTCTGATGGTACCCGCTACCCCTTCTGGACCTGGAAGAGTTCGGACAAATCTC 1447  
Db 1412 GTTCTCGGATGGATACCCATTCATCTTTGGACCTGGAAGAGTTCGACAGAACTCTC 1471  
QY 1448 CCAATTCACGAGGAAATCCAGAGGAGTTTGAATAAAAAGGCTTAACAGAAAGTGT 1507  
Db 1472 CCAACTGCACAGAGAGATCACAGAGGAGTCTGGAATAGAAGGCTACAGAAAGTGT 1531  
QY 1508 TTGCCAGTGGGTCTCGAGCAGTACTCTAAATATAGTAGGAGCCGGGAAGGCTTCTCG 1567  
Db 1532 TTGCACAGTGGGTCTCGAGCAGTACTCTAAACATGAGTGAAGCCATAAAGGCTTCTCG 1591  
QY 1568 TCAATGCTGTGCTACAAATCTATGGGCACGCTTTCGGAACCATCTTTTAAACTCAC 1627  
Db 1592 TCAAGTGTGTGCATACAAATTCCTTGGCACATCTTGTGAGACGATCTTTTAAACTCTC 1651  
QY 1628 CAGCCCTCTCCCTTTTCATCCAAAGACAACATCTCTCTTATCGGACCATTTGGGCTCTGC 1687  
Db 1652 CAGCCCTCTCCCTTTTCATCCAAAGACAACATCTCTCTTATGCAACAATTTGGTGTCTC 1711  
QY 1688 TCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAATTTAGT 1747  
Db 1712 TCCTCTTCATTTGCTTTTAAACCTGCTAAATTTGTCAAGTACAAAAGCAATTTAGT 1771  
QY 1748 ACGAGTCACTCGAGATGATCCAGGTGACTGGCCCTCGATACAGAGTACTTCTAGC 1807  
Db 1772 ATGAAGCCAGCTACAGATGTACAGGTGACCGGCTCTCGATATGAGTACTTCTAGC 1831  
QY 1808 TTGACTTCAGGAGTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAGAACTTAGACT 1867  
Db 1832 TTGATTTTCAGAGAATATGAATATGATCTCAATAGGAGTTTCCAGAGAGAAATTTAGAT 1891  
QY 1868 TTGGAAGGCTCTGGGCTCTCGGCTTCGGAAGGTTGATGAACGCCAGCGCTATAGCA 1927  
Db 1892 TTGGAAGGCTACTAGGATCAGTGTCTTTTGGAAAGTGTGAACCAACAGCTTATGGA 1951  
QY 1928 TTAGTAAACGGGAGTCTCAATTCAGGTGGGCTGAAGATGCTAAAAGAGAGTGAACA 1987  
Db 1952 TTAGCAAAACAGGAGTCTCAATTCAGGTGGCGTCAAAATGCTGAAAGAAAGCAGACA 2011  
QY 1988 GCTGTGAAAAGAGAGTCTCATGTCGGAGCTCAAAATGATGATGACCCAGCTGGACACCATG 2047  
Db 2012 GCTCTGAAGAGAGGACACTCATGTGAGAACTCAAGATGATGACCCAGCTGGGAAGCCAG 2071  
QY 2048 ACAACATCGTGAATCTGCTGGGGGATGCACACTGTCAGGGCAGTGTACTTTGATTTTG 2107  
Db 2072 AGAATATTTGAACCTGCTGGGGGCTGCACACTGTCAGGACCAATTTACTTTGATTTTG 2131  
QY 2108 AATATTTGCTATGCTGACCTCTCACTACCTAAGAAAGTAAAAGAGAGAGATTTTACA 2167  
Db 2132 AATACTGTGTATGCTGATCTCTCAACTATCTAAGAGTAAAAGAGAGAGATTTTACA 2191  
QY 2168 GGACATGCACAGAGATTTTAAAGGAACATAATTTCAAGTCTTACCTACTTTCCAGGCA 2227  
Db 2192 GGACTTGCACAGAGATTTTCAAGGAACACAAATTTCAAGTTTTTACCCACTTTCCCAATCAC 2251  
QY 2228 ATTCAAAATCCAGCATGCTGCTTCCAGAGAGTTCAGTTACACCCGCTTGGATCAGC 2287  
Db 2252 ATCCAAATCCAGCATGCTGCTTCAAGAGAGTTCAGATACACCCGACTCGGATCAAA 2311  
QY 2288 TCTCAGGTTCAATGGGAATTCATTCATCTGAAGATGAGATTTGAATATGAAGAACAG 2347  
Db 2312 TCTCAGGCTTCATGGGAATTCATTCATCTGAAGATGAAATTTGAATATGAAGAACAG 2371  
QY 2348 AGAGCTGGCAGAGAGAGAGGAGAGATTTCAACGCTGCTGACGTTTGAAGACCTTCT 2407  
Db 2372 AAAGGCTG-----GAAGAGAGGAGGAGTGAATGTGCTTACATTTGAAGATCTTCT 2425  
QY 2408 GCTTTGCTTCAAGTGGCCCAAGGCATGGAATTCCTGGAGTTCAAGTCTGCTGTGCCACA 2467

Db 2426 GCTTTCATATCAAGTTGCCAAAGGAATGAATTTCTGCAATTTAAAGTCGTGTTCACA 2485  
QY 2468 GAGACCTGGCAGCCAGGAATGTTGGTCAACCCAGGGAGGTGGTGAAGATCTGTGACT 2527  
Db 2486 GAGACCTGGCCGCCAGGACGTGCTGTCAACCCAGGGAAAGTGGTGAAGATATGTGACT 2545  
QY 2528 TTGGACTGGCCGAGACATCTCTGAGCGACTCCAGCTACGCTGCTGAGGGCAACGACGGC 2587  
Db 2546 TTGGATTTGGCTCGAGATCATGATGATTCACACTATGTTGTGAGGGCAATGCCGCTC 2605  
QY 2588 TCCCGGTGAAGTGGATGGCCAGCAGAGCTTTATTGAAGGATCTACACAATCAAGAGTG 2647  
Db 2606 TGCCGTGTAATATGGATGGCCCGCAAGGCTGTTTGAAGGCACTACACCAATTAAGAGTG 2665  
QY 2648 AGCTCTGGTCCCTACGGCATCTCTCTGGAGATATTTTCACTGGGTGGAACCTTACC 2707  
Db 2666 ATGCTGGTCATATGGAATATTACTGTGGGAATCTTCTCACTTGGTGTGAATCCTTACC 2725  
QY 2708 CTGGCAATTCCTGTCGACGCTAACTTCTATPAACTGATTCAGAGTGGATTAAATGGAGC 2767  
Db 2726 CTGGCAATTCGGTTGATGCTAACTTCTACAACTGATTCAAATGGATTAAATGGATC 2785  
QY 2768 ACCCAATTCATCCACAGAGGATATACTTTGTAATGCAATCTCTGGCTTTTGACT 2827  
Db 2786 AGCCATTTTATGCTACAGAAGAAATATACATTATATGCAATCTCTGCTGGCTTTTGACT 2845  
QY 2828 CAAGAAGGGCCATCTTCCCAACCTGACATCTATTTTATAGTGTGAGTGGCAGAGG 2887  
Db 2846 CAAGAAGGGCCATCTTCCCTAATTTGACTTTCGTTTTAGATGTGAGTGGCAGATG 2905  
QY 2888 CAGAAAGAG - - - - -CATGTATCAGAACTCATCATCATCTACC 2924  
Db 2906 CAGAAAGCGATGTATCAGATGTGGATGGCGGTGTTTCGGAATGCTCTACACCTACC 2965  
QY 2925 AAAACAGGGGCCCCCTCAGCAGAGAGGGGGTTCAGAGC - CAGTGCACACAGCGCCAGG 2983  
Db 2966 AAAACAGGGGACCTTTTACAGCAGAGAGATGGATTTGGGGCTACTCTCTCCGCGGCTCAGG 3025  
QY 2984 T-GAAGATTCACAGAAAGAGTTAGCGAGAGGCTTTGGACCCGCGC - - - - -ACCTAG 3038  
Db 3026 TCGAAGATTCGTAGAGAAACAATTTAGTTTAAAGGACTTCACTCCCTCCACCTATCCCTAA 3085  
QY 3039 CAGGCTGTAGCGCAGAGCCAGATTTAGCCTCGCCTCT - - - - -GAGGAAGCCCTCACAGCG 3096  
Db 3086 CAGGCTGTAGATTACCAACAGAGATTAATTTCACTACATAAAGAAATCTATTATCAAC 3145  
QY 3097 CGTGTCTCGCTGGACTTTTCTAGATGCTGTGCTCCATTAATC - - - - -CAAGTGA 3149  
Db 3146 TGCTGCTTCCACGAGACTTTTCTCTAGAACCGCTCTGCGTTTACTCTTTTCAAGGGA 3205  
QY 3150 CTTCTATAAATCAACCTCTCTCCACAGCGGGGAGAGCCCAATTAATGAGACTTTGTTGG 3209  
Db 3206 CTTTGTGTAATAATCAAAATCATCTCTGTCAAGGAGGAGGAGCTGATAATGAACCTTTATTG 3265  
QY 3210 TGAGCCGCGCTACCTTGGGGGCTTTTCCACGAGCTTGAGGGGAAAGCCATCTATCTGAA 3269  
Db 3266 GAGCATTGATCTGCATCCAGGCTTCTCAGGCGGCTTGAGTAATTTGTACCTGAAG 3325  
QY 3270 TATAGTATATCTTGTGTAATAGTGAACAAACCAACCCGTTTTTTTGTGTAAGGGAAGC 3329  
Db 3326 TACAGTATATCTTGTGTAATATACATAAACA - - - - -AGCATTTTGTGTAAGGGAAGC 3378  
QY 3330 TAAATATGATTTTAAATATCTATGTTTAAATATCTATGTAATCTTTTTCATCTATTTAG 3389  
Db 3379 TAAATATGATTTT - - - - -AAGTCTATGTTTAAATATATGTAATTTTTCAGCTATTTAG 3436  
QY 3390 TGAATATTTTATGGATGGAATATACTTCTACTCTAATAAATAAATAAATAAATAA 3449  
Db 3437 TGAATATTTTATGGGTGGGAATAAATTTCTACTACAGAAATAAATAAATAAATAA 3496  
QY 3450 AAAA 3453  
Db 3497 AAAA 3500

RESULT 9  
AR071704  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AR071704  
Sequence 3 from patent US 5912133.  
AR071704  
AR071704.1 GI:7222592

Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 3501)

AUTHORS

LEMISCHKA, I. R.

TITLE

JOURNAL

FEATURES

1. 3501

/organism="unknown"

BASE COUNT 1068 a 709 c 784 g 940 t

ORIGIN

Query Match

Best Local Similarity

Matches 2775; Conservative

64.5%; Score 2227.2; DB 6; Length 3501;

79.6%; Pred. No. 0;

Mismatches 653; Indels 56; Gaps 10;

QY 8 GGCTACCGCGCTCCGGAGGCCATCGGGCGTGTGGCGCAGCGCAGCGCGGCTGC 67

Db 35 GGGGACCCCGGCTCCGGAGGCCATCGGGCGTGTGGC - - - - -CGCGACGCGGGCACCGTGC 91

QY 68 TGCTGCTGTTGTTTGTCTAGTAATGATTTCTGAGACCGTTTACAAACCAAGCACTGCTG 127

Db 92 CGCTGCTGTTGTTTCTGCAATGATATTGGACTATTACAAATCAAGATCGCTGCTG 151

QY 128 TGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGCTCATCAGCGGGAAAGCCATCAT 187

Db 152 TGATCAAGTGTGTTTAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 211

QY 188 CGTACCAGATGTCGAGGATCCCGAAGACCTCCAGTGTACCCGAGGCGCCAGAGTG 247

Db 212 CATATCCATGGTATCAGAAATCCCGAAGACCTCGGGTGTGCTTGAGACCCAGAGCT 271

QY 248 AAGGGACGTTATATGAAGCGCCACCGTGGAGTGGCGGAGTCTGGGTTCATACCCCTGC 307

Db 272 CAGGGACAGTGTACGAGCTGCCGCTGTGGAAGTGTATCTGCTTCCATCACATGC 331

QY 308 AAGTGCAGCTGCCACCCAGGGACCTTTCCCTGCTGCTGGTCTTTAAGCACAGCTCCC 367

Db 332 AAGTGCCTGTCGATGCCCGAGGAACATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 391

QY 368 TGGGCTGCCAGCGCCTTTGATTTACAAACAGAGGAATCGTTTCCATGCCCATCTTGA 427

Db 392 TGAATGCCAGCCACATTTTGTATTTACAAACAGAGGAGTGTGTTTCCATGGTCTCATTTGA 451

QY 428 ACCTGACAGAGACCCAGGAGAAATACCTACTTCCATATTTCAGAGGAAAGCCGCAACT 487

Db 452 AATGACAGAAACCCAGCTGGAGATACCTACTTTTATTCAGAGTGAAGCTACCAATT 511

QY 488 ACACAGTACTGTTTACAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGAGACCTT 547

Db 512 ACACATATTTTACAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGAGACCTT 571

QY 548 ACTTTAGAGAGATGGAAACCCAGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 607

Db 572 ACTTTAGAGAGATGGAAACCCAGAGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 631

QY 608 CCACTGTGGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 667

Db 632 CGATCGTGGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 691

QY 668 CTGTTGTCAGAAAGGAGGAAAGTACTTTCATGATGTTGTTGCGGAAACACATCATGATGCT 727

Db 692 CTGTTGTCAGAAAGGAGGAAAGTACTTTCATGATGTTGTTGCGGAAACACATCATGATGCT 751



Qy 728 GTGCTAGAAATGCACTGGCGCGCAATGCACCAAGCTGTTCCACATAGATCTAAACAGG 787  
Db 752 GTGCCAGAAATGAACCTGGCGAGGAATGCACACAGCTGTTCCAAATAGATCTAAATCAAA 811  
Qy 788 CTCCTCAGAGCACACTGCCCCAGTATTCTCTGAAAGTGGGGAAACCCCTTGCGATCAGGT 847  
Db 812 CTCCTCAGAGCACATGCCCACAAATATTCTCTAAAGTAGGGAAACCCCTTATGGTAAGGT 871  
Qy 848 GTAAGGCCATCCATGTGAACCATGGATTGGGGCTCACCTGGGAGCTGGAAGACAAAGCCC 907  
Db 872 GCAAAGCTGTTCATGTGAACCATGGATTGGGGCTCACCTGGGGAATTAGAAACCAAGCAC 931  
Qy 908 TGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATGATTCGGA 967  
Db 932 TCGAGGAGGCGAACAATACTTTGAGATGAGTACCTATTCAACAAACAGAACTATGATACCGA 991  
Qy 968 TTCTCTTTGGCCCTTTGTCTTCCGTGGGAAGGAACGACACCCGGATATTACACCTGCTCTT 1027  
Db 992 TTCTGTTTGTGTTATCATCATGTCAGTGGCAAGAACGACACCCGGATACACTTGTTCCT 1051  
Qy 1028 CTTCAAGACCCAGCCAGTCAGCTGAGTGGTGACCATCTAGAAAGGGTTTATAACG 1087  
Db 1052 CTTCAAGACATCCAGTCAATCAGCTTTGGTTACCATCGTAGGAAGGGATTTATAAATG 1111  
Qy 1088 CTACCACTCGCAAGAAAGAGTATGAATTTGACCCGTAGCAAAAGTTCCTTCAGTCA 1147  
Db 1112 CTACCAATCAAGTGAAGATTATGAATTTGACCAATATGAAGAGTTTGTCTCTGTCA 1171  
Qy 1148 GGTTTAAAGGTCACCCAGATCCGATCGACGTGGATCTCTCTCAAGCCCTCATTTCCCT 1207  
Db 1172 GGTTTAAAGCTACCCACAATCAGATCTACGTGGACCTTCTCTCGAAATCAITTCCT 1231  
Qy 1208 GTGAACAGAGGCGCTGGAGATGGGTACACATATCTAAATTTTGGCATCATAGAACA 1267  
Db 1232 GTGAGCAAAAGGGTCTTGATAACGGATACACATATCCAAAGTTTGGCAATCATAGCAC 1291  
Qy 1268 AGCCAGGAGTACATATCTCTATCAGAAATGATGACCCGATGACCAAAATGTTCA 1327  
Db 1292 AGCCAGGAGATATATTTCCATCGAANAATGATGATGCCCAATTTACCAAAATGTTCA 1351  
Qy 1328 CGTGAATATAAGAAAGAACTCAAGTCTAGCAAAATGCTCTAGCCAGCCAGCGCTCT 1387  
Db 1352 CGCTGAATATAAGAAAGAACTCAAGTCTCGCAGAGCATCGCAAGTCAGCGCTCT 1411  
Qy 1388 GTTCCTCTGATGGCTACCCGCTACCTCTTGAGCTTGGAAAGTGTTCGGACAAATCTC 1447  
Db 1412 GTTCTCGGATGATACCCATTACCATCTTGGACCTTGGAAAGTGTTCAGACACAGTCTC 1471  
Qy 1448 CCAATTCACGAGGAATCCAGAGAGTGTGGAATATAAGGCTTAACAGAAAGTGT 1507  
Db 1472 CCAACTGCAGAAAGAGATCAGAAAGAGTCTGGAATAGAAGGCTAACAGAAAGTGT 1531  
Qy 1508 TTGGCCAGTGGGTGTGAGCAGTACTCTAAATATAGTGAAGCCGGGAAAGGCTTCTGG 1567  
Db 1532 TTGGACAGTGGGTGTGAGCAGTACTCTAAACATGAGTGAAGCCATAAAGGTTCCCTGG 1591  
Qy 1568 TCAATGCTGTGCGTACAAATCTATGGCAGCTTGGCAACCATCTTTTAAACTCAC 1627  
Db 1592 TCAAGTGTGTGCATACAAATTTCCCTTGGCATCTTGTGAGACGATCTTTTAAACTCTC 1651  
Qy 1628 CAGGCCCTTCCCTTTCATCCAGACACATCTCTCTATGCGACCATTTGGGCTCTGTC 1687  
Db 1652 CAGGCCCTTCCCTTTCATCCAGACACATCTCTCTATGCGAACAATTTGTGTTGTC 1711  
Qy 1688 TCCCTCTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAAGCAATTTAGGT 1747  
Db 1712 TCCTCTCATTTGCTTTTACCCCTGCTAAATTTGTCACAGTACAAAAGCAATTTAGT 1771  
Qy 1748 ACAGAGTCTGATGCAGATGATCCAGGTGATGCGCCCTCGATTAACGAGTACTTCTACG 1807  
Db 1772 ATGAAGCCAGCTACAGATGATACAGGTGACCGGCTCTCTCAGATAATGATGACTTCTACG 1831

Qy 1808 TTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCGCCAGAGAGAACTTAGAGT 1867  
Db 1832 TTGATTTCAGAGAAATGAATATGATCTCAATGGGAGTTTCCAAGAGAAATTTAGAGT 1891  
Qy 1868 TTGGAAGGCTCTGGGCTTGGCGCTTCCGGAGGCGTGATGAACGCCACGGCTATAGGCA 1927  
Db 1892 TTGGAAGGCTACTAGGATCAGGTCTTTTGGAAAACTGATGAACGCAACAGCTTATGGAA 1951  
Qy 1928 TTAGTAAACGGAGCTCAATTCAGTGGCGGTGAAGATGCTTAAAGAGAAAGCTGACA 1987  
Db 1952 TTAGCAAAACAGAGTCTCAATCCAGGTTGCCGCTCAAAATGCTGAAGAAAGAACAGACA 2011  
Qy 1988 GCTGTGAAAAAGAGCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047  
Db 2012 GCTCTGAAGAGAGGCACTCATGTGAGAACTCAAGATGATGACCCAGCTGGGAAGCCAG 2071  
Qy 2048 ACAACATCGTGAATCTGTGGGGGCAATGACACTGTGAGGGCCAGTGACTGATTTTTTG 2107  
Db 2072 AGAATATTGTGAACCTGCTGGGGGGTGACACTGTGAGGACCAATTTACTTGAATTTTG 2131  
Qy 2108 RATATTGTGTATGTGTGACCTCTCAACTACTTAAGAAGTAAAGAGAGAAAGTTTCA 2167  
Db 2132 AATAGTGTGTATGTGTGATCTCAACTATCTAAGAAGTAAAGAGAGAAATTTCA 2191  
Qy 2168 GGACATGACAGAGATTTTAAAGGAACATAATTTCAAGTTCTTACCCCTACTTTCCAGG 2227  
Db 2192 GGACTTGGACAGAGATTTCAAGGAACACAAATTTCAAGTTTACCCCACTTTTCCAATC 2251  
Qy 2228 ATTCAAATTCAGATGCTGTGTTACAGAGAGTTTCAAGTTTACCCGCCCTTGGATCAGC 2287  
Db 2252 ATTCAAATTCAGATGCTGTGTTCAAGAGAGTTTCAAGATACACCCGGACTCGGATCAA 2311  
Qy 2288 TCTCAGGTTCAATGGGAATTCATTCATCTCAAGATGAGATGAAATGAAATGAAACCA 2347  
Db 2312 TCTCAGGTTTCAATGGGAATTCATTCATCTCAAGATGAGATGAAATGAAATGAAACCA 2371  
Qy 2348 AGAGCTGTCAGAAAGAGAGAGGAGATTTGAACGTGCTGAGCTTTGAAGACCTCTCTT 2407  
Db 2372 AAAGCTG-----GAAGAGAGGAGGACTTGAATGCTGTACATTTGAAGATCTCTT 2425  
Qy 2408 GCTTGGCTTACCAAGTGGCCAAAGCATGGAATTCCTGGAGTTCAAGTCTGTGTCCACA 2467  
Db 2426 GCTTGGCATATCAAGTTGGCCAAAGGAATGGAATTTCTGGAATTTAAGTCTGTGTCCACA 2485  
Qy 2468 GAGACCTGCGACGAGGAATGCTGTCACCCAGGAGGAGTGGTGAAGATCTGTGACT 2527  
Db 2486 GAGACCTGCGCCGAGGAGGAGTGTGTCACCCAGGAGGAGTGGTGAAGATATGTGACT 2545  
Qy 2528 TTGGAGTGGCCGAGACATCTCTGAGGACTTCCAGCTACGTCGAGGGGCAACGACGCG 2587  
Db 2546 TTGGATGCTCGAGATATCATGAGTATTCCAACTATGTTGTGAGGGCAATGCCGCTC 2605  
Qy 2588 TCCCGGTGAAGTGGATGGCACCCGAGAGCTTTATTTGAAGGATCTACAAATCAAGAGT 2647  
Db 2606 TGCTGTAAATGGATGGCCCGGAAAGCTGTTTGAAGGCACTACACATTAAGAGT 2665  
Qy 2648 AGCTGTGGCTTACGCACTCTCTGAGAGATATTTTCACTGGGTGTGAACCTTTACC 2707  
Db 2666 ATGTCTGTGATGAATATTTACTGTGGAAATCTTCACTTGGTGTGAATCTTTACC 2725  
Qy 2708 CTGGCATTCCTGTGACGCTAACTCTTATAAATGATTCAGAGTGGATTTAAATGGAGC 2767  
Db 2726 CTGGCATTCGGTGTGATGCTACTTCTACAACTGATTCAAATGGATTTAAATGGATC 2785  
Qy 2768 AGCCATTTATGCCAGAGAGGATATCTTTGTAATGCAATCCTGCTGGCTTTTGAAT 2827  
Db 2786 AGCCATTTATGCTACAGAGAAATATACATTAATGCAATCCTGCTGGCTTTTGAAT 2845  
Qy 2828 CAAGAGAGGCGCATCTTCCCAACCTGACTTCATTTTAGGATGTCAGCTGGCAGAG 2887  
Db 2846 CAAGAGAGGCGCATCTTCCCTTAATTTGACTTCGTTTTTAGGATGTCAGCTGGCAGAG 2905  
Qy 2888 CAGAAGAAG-----CATGTATCAGAACAATCCATCTTACC 2924



Db 1172 GGTTTAAAGCCTACCCACAAATCAGATGATCGTGGACCTTCTCTCGAAAAATCAATTCCTT 1231  
Qy 1208 GTGACAGAGGCGCTGGAGATGGGTACACATATCTAAATTTGGCATCATAGACAA 1267  
Db 1232 GTGAGCAAAAGGGTCTTGATTAACGGATACACATATCCAAATTTTGGCAATCATAGCAC 1291  
Qy 1268 AGCCAGGAGATACATATTTCTATCAGAAAAATGATGACGCCAGTTTACCAAAAATGTTCA 1327  
Db 1292 AGCCAGGAGATATATTTCCATCGACAGAAATGATGCCCCAATTTACCAAAAATGTTCA 1351  
Qy 1328 CGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAGCGGCTCT 1387  
Db 1352 CGCTGAATATAAGAAAGAAACCTCAAGTGTGCGAGAAAGCATCGGCAAAAGTTCAGGCGTCT 1411  
Qy 1388 GTTCCTGTGAGCTACCCGTACCTCTTGGACCTGGAAGAGTGTTCGACCAAAATCTC 1447  
Db 1412 GTTCTCGGATGAGTACCCATCTTGGACCTGGAAGAGTGTTCGACCAAAATCTC 1471  
Qy 1448 CCAATTGCACGAGGAAATCCAGAAAGAGTGTTCGAAATAAAAGGCTAACAGAAAAGTGT 1507  
Db 1472 CCAACTGCACAGAGAGATACAGAGAGGCTGGAATAGAAAGGCTAACAGAAAAGTGT 1531  
Qy 1508 TTGGCCAGTGGGTGTGAGCAGTACTCTAAATATAGTGTAGCGCGGAAAGGCTTCTGG 1567  
Db 1532 TTGGACAGTGGGTGTGAGCAGTACTCTAAACATAGTGAAGCCATAAAAGGTTTCCTGG 1591  
Qy 1568 TCAATGCTGTGCTACAAATCTATGGCCAGCTTTCGGAACCAATCTTTTAAACTCAC 1627  
Db 1592 TCAAGTGTGTGCTACAAATCTTCCCTTGGCAGCATCTTGTGAGCAGTCTTTTAAACTCTC 1651  
Qy 1628 CAGGCCCTTCCCTTTATCCAAGACAACTCTCTCTTATGCGACCAATTTGGGCTGTCTC 1687  
Db 1652 CAGGCCCTTCCCTTTATCCAAGACAACTCTCTCTTATGCGACCAATTTGGTGTCTC 1711  
Qy 1688 TCCCTTCATTTGTTCTCATTTGTGTTGATCTGCGCAAAATACAAAAAGCAATTTAGGT 1747  
Db 1712 TCCCTTCATTTGTTTAAACCTGCTAAATTTGTCAAGTGTACAAAAAGCAATTTAGGT 1771  
Qy 1748 ACGAGATCAGCTCAGATGATCCAGGTGACTGGCCCTCGATTAACAGTACTTCTAGC 1807  
Db 1772 ATGAAGCCAGCTCAGATGATGAGTGTGAGCGGCTCTCGATTAATGAGTACTTCTAGC 1831  
Qy 1808 TTGACTTCAGGAGCTATGAATATGACCTTAAGTGGGAGTTCGCGAGAGAGAACTTTAGAT 1867  
Db 1832 TTGATTCAGAGATATGAATATGATCTCAATGGAGTTTCCAGAGAGAAATTTAGAT 1891  
Qy 1868 TTGGAGAGTCTCTGGGCTCTGGGCTTTGGGAGGTTGATGAACCCACGCGCTATGGCA 1927  
Db 1892 TTGGAGAGTACTAGGATCAGTGTGCTTTTGGAAAAAGTGTAAACGCAACAGCTTATGGAA 1951  
Qy 1928 TTAGTAAACGGGAGTCTCAATTCAGTGGCGGTCAAGATGCTAAAGAGAAAGCTGACA 1987  
Db 1952 TTAGCAAAACAGGAGTCTCAATTCAGTGGCGGTCAAAATGCTGAAGAAAGAAAGCAGACA 2011  
Qy 1988 GCTGTAAAAAGAAAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047  
Db 2012 GCTGTAAAGAGAGGACATCATGTGAGAACTCAAGATGATGACCCAGCTGGGAAGCCAG 2071  
Qy 2048 ACAACATCGTGAATCTGTGGGGGATGCACACTGTGAGGGCAGTGTACTTTGATTTTG 2107  
Db 2072 AGAATATTGTGAACCTGTGGGGGCTGCACACTGTGAGGACCAATTTACTTTGATTTTG 2131  
Qy 2108 AATATTGTGTATGGTGACCTCACTACCTAAGAAAGTAAAGAGAGAAAGTTCACA 2167  
Db 2132 AATACTGTGTATGGTATCTCTCAACTATCTAAGAGTAAAGAGAGAAATTTTACA 2191  
Qy 2168 GGACATGACAGAGATTTTAAAGAACTAATTTCACTTCTTACCTTCTTCCAGGCAC 2227  
Db 2192 GGACTTGGACAGAGATTTTAAAGAACTAATTTCACTTCTTACCTTCTTCCATCAC 2251  
Qy 2228 ATTCAAATTCAGATGCTGTGTCAGAGAAAGTTCAGTTACACCGCGCTTGGATCAGC 2287  
Db 2252 ATCCAAATTCAGATGCTGTGTCAGAGAAAGTTCAGATACACCGGACTCGGATCAA 2311

Qy 2288 TCTCAGGTTCAATGGAAATTCATTCATCTGAAGATGAGATTTGAATATGAATAACAGA 2347  
Db 2312 TCTCAGGCTTCATGGGAATTCATTCATCTGAAGATGAGATTTGAATATGAATAACAAA 2371  
Qy 2348 AGAGCTGGCAGAGAAAGAGAGAGATTTGAACGTGTGAGCTTTGAAGACCTCTCTT 2407  
Db 2372 AAAGCTG-----GAAGAGAGGAGGACTTGAATGTCTTACATTTGAAGATCTCTT 2425  
Qy 2408 GCTTTGCTACCAAGTGGCCAAAGCATGGAATTCCTGGAGTCAAGTCTGCTGTCCACA 2467  
Db 2426 GCTTTGCTATCAAGTGGCCAAAGGAATGGAATTTCTGGAATTTAGTCTGTGTTTACA 2485  
Qy 2458 GAGACCTGGCAGCAGGAATGTGTGTCACCCAGGAGGTGTGAAGATCTGTGACT 2527  
Db 2486 GAGACCTGGCCCGCAGGACGTGTGTCACCCAGGAAAGTGTGAGATATGTGACT 2545  
Qy 2528 TTGAGCTGGCCCGCAGACATCTGAGGACTTCCAGCTACGTCGTCAGGGGCAACGACGG 2587  
Db 2546 TTGGATTTGGCTCGAGATATCATGAGTGAATTCCAACTATGTTGTGAGGGCAATGCCGTC 2605  
Qy 2588 TGCCCGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACACAAATCAAGAGT 2647  
Db 2606 TGCCCTGTAATGGAATGGGCCCCGAAAGCTGTTTGAAGCATCTACACATTAAGAGT 2665  
Qy 2648 AGTCTGGTCTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAACCTTTACC 2707  
Db 2666 ATGCTGTGCTATATGGAATTTACTGTGGAATCTTCTCACTTTGTTGTAATCTTACC 2725  
Qy 2708 CTGGCATTCCTGTCAGCCTAACTTCTATAAATGATTCAGAGTGAATTTAAATGGAGC 2767  
Db 2726 CTGGCATTCGGTGTGATGCTAACTTCTACAAATGATTTCAAAATGGATTTAAATGGATC 2785  
Qy 2768 ACCCATCTATGCCAGAGGATATACTTTGTAATGAATCCTGCTGGGCTTTTACCT 2827  
Db 2786 AGCCATTTTATGCTACAGAAAGATATACATTAATGCAATCCTGCTGGGCTTTTACCT 2845  
Qy 2828 CAAGAGCGGCTCTTCCCAAGCTGACTTCTATTTAGGATGTGACCTGGCAGAG 2887  
Db 2846 CAAGAGCGGCTCTTCCCAAGCTGACTTCTATTTAGGATGTGACCTGGCAGAG 2905  
Qy 2888 CAGAAGAG-----CATGTATCAGAACATCTCATCTACCTTACC 2924  
Db 2906 CAGAAGAGGATGTATCAGAAATGTGATGGCGCTGTTCCGAATGTCTTCAACCTTACC 2965  
Qy 2925 AAAACAGCGGCTCTCAGCAGAGAGCGGCTCAGAGC--CCAGTCCGACAGCGCCAG 2983  
Db 2966 AAAACAGCGGCTCTTCCAGCAGAGAGATGGATTTGGGCTTACTCTCTCCGAGGCTCAG 3025  
Qy 2984 T--GAAGATTACAGAGAAAGATTTAGCGAGGAGGCTTTGGACCCCGCC-----ACCCTAG 3038  
Db 3026 TCGAAGATTCGTAGAGAACATTTAGTTTAAAGACTTCTATCCCTCCACCTATCCCTAA 3085  
Qy 3039 CAGGCTGTAGACCGCAGAGCCCAAGATTAGCTTCGCTCT--GAGGAAGCGCCCTACAGC 3096  
Db 3086 CAGGCTGTAGATTACCAAAAGATTAATTTTATCATAAAAGAAATCTATATCAAC 3145  
Qy 3097 GGTGCTTCCGCTGGACTTTTCTAGATGCTGTCTGCCATTACTC-----CAAGTGA 3149  
Db 3146 TGCTGCTTCCAGCAGCTTTTCTAGAGCGCTCTGCTTACTTCTTTTCAAGGGA 3205  
Qy 3150 CTTCATAAAATCAAACTCTCTCTGACAGCGGAGAGCCAAATATAGACTTTGTGG 3209  
Db 3206 CTTTGTAAATCAAAATCATCTCTGACAGGAGGAGGAGCTGATTAATGAATTTATTG 3265  
Qy 3210 TGAGCCCGCTCACCTGGGCGCTTTCCAGAGCTTGAGGGGAAAGCCATGTATCTGAAA 3269  
Db 3266 GAGCATTTGATCTGCATCCAGGCTTCTCAGGCGGCTTGAAGTGAATTTGCTACCTGAAG 3325  
Qy 3270 TATAGTATATCTTTGTAATATGTAACCAAAACCCGCTTTTTCGTTAAGGAAGC 3329  
Db 3326 TACAGTATATCTTGTGTAATATACATAAAACAAA-----AGCATTTTTCGTAAGGAGAGC 3378



[illegible]

Db	2726	CTGCATTCGGTTGATGCTAACTTCTACAAACTGATCAAAATGGATTTAAATGGATC	2787
Qy	2768	AGCCATCTTATGCCACAGAAGGATATACATTTGTAAATGCAATCTGCTGGGCTTTTGACT	2827
Db	2786	AGCCATTTTATGCTACAGAAGAATATACATTAATGCAATCTGCTGGCTTTTGACT	2845
Qy	2828	CAAGGAAGCGCCATCCTTCCCAACCTGACTTCATTTTATAGGATGTCACCTGGCAGAGG	2887
Db	2846	CAAGGAACGCCATCCTTCCCTAAATTGACTTCGTTTTAGGATGTCAGCTGGCAGATG	2905
Qy	2888	CAGAAAGAG-----CATGTATCAGAACTCCATCCATCTACCTACC	2924
Db	2906	CAGAAGAAGCGATGATCAGAAATGTGATGGCGGTTCGGAATGTCTTCACACCTTACC	2965
Qy	2925	AAAACAGCGCGCCCTCAGCAGAGAGCGGGCTCAGAGC--CCAGTCCGCCACAGCGCCAGG	2983
Db	2966	AAAACAGCGGACCTTTCAGCAGAGAGATGATTTGGGGCTACTCTCTCCCGCAGGCTCAGG	3025
Qy	2984	T-GAAGATTCACAGAGAAAGAAGTTACGAGGAGGCGCTTGAGCCGCC-----ACCCTAG	3038
Db	3026	TCGAAGATTCGTAGAGGAACAATTTAGTTTTAAGGACTTCATCCCTCCACTATCCGTAA	3085
Qy	3039	CAGGCTCTAGACGCGACAGGCGAAGATTAGCTCGCCTCT--GAGGAAGCGCCCTACAGCG	3096
Db	3086	CAGGCTGTAGATTACCAAAAACAAGATTAAATTCATCACTAAAAGAAATCTATTATCAAC	3145
Qy	3097	CGTTGCTTCGTGGACTTTTCTCTAGATGCTGTGCTCATTACTC-----CAAAGTGA	3149
Db	3146	TGCTGCTTACCAGACTTTTCTCTAGAGCGCTCTGCTTTACTCTTTTCAAGGGA	3205
Qy	3150	CTTCTATAAATCAAACCTCTCTCGCACAGCGGGAGAGCCAAATAATAGACTGTGTGG	3209
Db	3206	CTTTTGTAAAAATCAATCATCTCTGCACAGGCGAGGAGGAGCTGATATGAACCTTATTG	3265
Qy	3210	TGACCCCGCTACCTGGGGCCCTTTCCACGAGCTTGAGGGGAAGCCATGCTATCTGAA	3269
Db	3266	GAGCATTTAGTGTGCATCCAGGCGCTTCTCAGCGCGGCTTGAGTGAATGTGTACTCTGAAG	3325
Qy	3270	TATAGTATATCTTGTAAATACGTGAACAAACCAACCCCGCTTTTTCGCTAAGGGAAGC	3329
Db	3326	TACAGTATATCTTGTAAATACATAAACA-----AGCATTTTTCGTAAGGAGAAGC	3378
Qy	3330	TAAATATGATTTTAAAAATCATGTATTTTAAATACATGATGTAATCTTATTTAG	3389
Db	3379	TAATATGATTTT--AAGTCTATGTTTAAATATAATGTAATAATTTTTCAGCTATTTAG	3436
Qy	3390	TGATATATTTATCGATGGAATAAATCTTCTACTGTAAATAAAAAAAAAAAAAAAAAA	3449
Db	3437	TGATATATTTATGGTGGGAATAAATTTCTACTACAGAAAAAAAAAAAAAAAAAAAA	3496
Qy	3450	AAAA 3453	
Db	3497	AAAA 3500	
RESULT 12			
LOCUS	I40109	3476 bp	DNA
DEFINITION	Sequence 1 from patent US 5618709.	linear	PAT 13-MAY-1997
ACCESSION	I40109		
VERSION	I40109.1	GI:2083114	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 3476)		
TITLE	Gewirtz,A.M., Smali,D. and Civin,C.I.		
JOURNAL	Antisense oligonucleotides specific for STK-1 and method for		
FEATURES	inhibiting expression of the STK-1 protein		
source	Patent: US 5618709-A 1 08-APR-1997;		
	Location/Qualifiers		
	1..3476		
	/organism="unknown"		

BASE COUNT	1044 a	709 c	783 g	940 t	
ORIGIN					
Query Match	63.9%; Score 2204.8; DB 6; Length 3476;				
Best Local Similarity	79.5%; Pred. No. 0;				
Matches 2752; Conservative	0; Mismatches 652; Indels 56; Gaps 10;				
Qy	8	GGCTACCGCGCTCCCGAGGCCATCGGGGCTTGGCGCAGCGAGCGCGCGGCTGC	67		
Db	35	GGGACCGCGGCTCCCGAGGCCATCGCGGCTTGGC---GCGGACGCGCGCACCGTGC	91		
Qy	68	TGCTGCTGTGTTTGTTCAGTAATGATCTCTGAGACGCTTACAAACCAAGACCTGCTG	127		
Db	92	CGCTGCTGTGTTTGTTCAGTAATGATCTCTGAGACGCTTACAAACCAAGACCTGCTG	151		
Qy	128	TGATCAAGTGTGTTTAAATCAGTCATGAGACAATGGCTCATACGCGGAAGACCATCAT	187		
Db	152	TGATCAAGTGTGTTTAAATCAGTCATGAGACAATGGCTCATACGCGGAAGACCATCAT	211		
Qy	188	CGTACCGAATGGTCGGAGGATCCCGAGAACCTCCAGTGACCCGAGGCGCGAGAGTG	247		
Db	212	CATATCCCATGGTATCAGATCCCGAGAACCTCCAGTGACCCGAGGCGCGAGAGTG	271		
Qy	248	AAGGAGCGGTATATGAAGCGGCCACCGTGGAGTGGCGAGTCTGGTCCATCACCCCTGC	307		
Db	272	CAGGACAGTGTACGAGCTGCCGCTGTGGAAGTGGATGATCTGCTTCCATCACACTGC	331		
Qy	308	AAGTGCAGCTGCGCACCCCGAGGAGCTTCCCTGCTCTGGCTCTTTAAGCACAGCTCC	367		
Db	332	AAGTGCCTGGTGCATGCCCGCAGGGAACATTCCTGCTCTGGGCTCTTTAAGCACAGCTCC	391		
Qy	368	TGGGCTGCCAGCGCCTTCATTTACAAACACAGGAATGCTTTCCATGGCCATCTTGA	427		
Db	392	TGAATGCCAGCCACATTTGATTTACAAACACAGGAGTGTTCATGGTGCATTTTGA	451		
Qy	428	ACGTGCACAGACCCAGGAGGAGAAATACCTACTCCATTCATCAGAGCGGAAGCGCCACT	487		
Db	452	AAATGCACAGAACCCAGCTGGAGATACCTACTTTTATTCAGAGTGAAGCTACCAAT	511		
Qy	488	ACACAGTACTGTTACAGTGAATGTAAGATATACACAGCTGTATGTGCTAAGGAGACCTT	547		
Db	512	ACACAATATGTTTACAGTGAATGTAAGAAATACCTGCTTTTACACATTAAGAAGACCTT	571		
Qy	548	ACTTTAGAAATGAAACACAGGATGCATGCTCTGCTGATCTCCAGAGGTGTTCGGAGC	607		
Db	572	ACTTTAGAAATGAAACACAGGAGCCCTGGTCTGCATATCTCAGAGCTGTTCAGAGC	631		
Qy	608	CCACTGTGGAGTGGTGTCTGCAGCTCCCGAGGAAAGCTGTAAAGAAAGGCCCTG	667		
Db	632	CGATCGTGAATGGTGTCTTGGGATTCACAGGCGGAAAGCTGTAAAGAAAGTCCAG	691		
Qy	668	CTGTGTGTAAGAGGAGAAAGTACTTTCATGATGTTGTCGGAACACAGATCAGATGCT	727		
Db	692	CTGTGTGTTAAAGAGGAGAAAGTCTCATGATATTTTGGGACGGAATGATGATGCT	751		
Qy	728	GTGCTAGAAATGCTGCGCCGATGCACCAAGCTGTTACCATAGATCTAAACCCAGG	787		
Db	752	GTGCGAGAAATGAACTGGGCGGGAATGCACAGGCTGTTCAATAGATCTAAATCAA	811		
Qy	788	CTCCTCAGAGCACACTGCGCCAGTATTCTCTGAAAGTGGGGAACCTTGTGGATCAGGT	847		
Db	812	CTCCTCAGACCACATGCGCAATATTCTTCTTAAGTAGGGAACCTTTATGATGAAGT	871		
Qy	848	GTAAAGCCCATCATGTGAACCATGGATTCGGGCTCACCTGGAGCTGGGAAGCAAGCCC	907		
Db	872	GCAAGCTGTTTCATGTGAACCATGGATTCGGGCTCACCTGGGAATTTAGAAAACAAAGCAC	931		
Qy	908	TGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATGATCGGA	967		
Db	932	TCGAGGAGGCGCACTACTTTGAGATGAGTACCTATTTCACAAACAGAACATGATACGGA	991		
Qy	968	TTCTCTTGGCCCTTGTGCTCCGTGGGAGGAACGACACCCGATATTACCTGCTCTT	1027		
Db					
Db	992	TTCTGTTTGGCTTTTGTATCATCATAGTGGCAAGAAACGACACCCGGATFACATCTTGTCT	1051		
Qy	1028	CCTCAAGACCCCGAGCCAGTCAGGTTGGTGACCATCTCTAGAAAAGGTTTATAAAGC	1087		
Db	1052	CTTCAAGACATCCCGAGTCATCAGCTTTGGTTTACCATCTGTAGAAAAGGATTTATAAATG	1111		
Qy	1088	CTACAGCTCGCAAGAAGAGTATGAAATTTGACCCGTAGCAAAAGTTCTGCTCTCAGCTCA	1147		
Db	1112	CTACCAATTCAGATGAAGATATGAAATTTGACCAATATGAAGAGTTTCTGTTTCTGTCA	1171		
Qy	1148	GGTTAAAGCGTACCCAGCAATCCAGTCAGCTGATCTTCTCTCAGGCTCATTTCCCT	1207		
Db	1172	GGTTAAAGCCTTACCCACAAATCAGATGTATGACCTTCTCTCGAAATCATTTCCCT	1231		
Qy	1208	CTGAACAGAGGCGCTGAGGATGGTACAGCATATCTAAATTTTGGCGATCATFAAGACA	1267		
Db	1232	GTGAGCAAAAGGCTCTGATACGGATACAGATATCTCAAGTTTTCATCATAGCACC	1291		
Qy	1268	AGCCAGGAGTACATATTTCTATGCAGAAAATGATGACCCCGAGTTACCAAAATTTTCA	1327		
Db	1292	AGCCAGGAGATATATATTTCCATGCAGAAAATGATGATGCCCAATTTACCAAAATTTCA	1351		
Qy	1328	CGCTGAATATAGAAGAAACCTCAAGTGTAGCAATGCCTCAGCCAGCCGCTCCT	1387		
Db	1352	CGCTGAATATAGAAGAAACCTCAAGTGTAGCAATGCCTCAGCAAGTCAGGCGCTCCT	1411		
Qy	1388	GTTCTCTGTAGGCTACCGCTACCTCTTGGACCTGGAAGAGTCTTCGGAACAATCTC	1447		
Db	1412	GTTCTCTGTAGGCTACCGCTACCTCTTGGACCTGGAAGAGTCTTCGGAACAATCTC	1471		
Qy	1448	CCAATTCACCGGAGGAAATCCCGAAGAGTGTGGAAATAAAAGGTAAACAGAAAAGTGT	1507		
Db	1472	CCAATTCACCGGAGGAAATCCCGAAGAGTGTGGAAATAAAAGGTAAACAGAAAAGTGT	1531		
Qy	1508	TTGGCCAGTGGTGTGCGACACTACTTAATATAGTGTAGGCGCGGGAAGGCTCTCG	1567		
Db	1532	TTGGACAGTGGTGTGCGACACTACTTAATATAGTGTAGGCGCGGGAAGGCTCTCG	1591		
Qy	1568	TCAAATGCTGCGTACAAATCTATGGCGAGTCTTTCGAAACCATCTTTTAAACTCAC	1627		
Db	1592	TCAAATGCTGCGTACAAATCTTTCGCGACATCTTGTGAGAGGATCTTTTAAACTCAC	1651		
Qy	1628	CAGGCGCTTCCCTTTCATCCAGACACATCTCTTCTATGCGACCATTTGGGCTCTGTC	1687		
Db	1652	CAGGCGCTTCCCTTTCATCCAGACACATCTCTTCTATGCAACAATTTGGTGTGTC	1711		
Qy	1688	TCCCTCTCAATGTTGTTCTCATGTTGATCGCCACAAATACAAAAGCAATTTAGGT	1747		
Db	1712	TCCCTCTCAATGTTGTTCTCATGTTGATCGCCACAAATACAAAAGCAATTTAGGT	1771		
Qy	1748	ACGAGAGTCAGCTGCAGATGATCCAGTGACTGCGCCCTTGGATACGAGTACTTCTACG	1807		
Db	1772	ATGAAGCCAGCTGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1831		
Qy	1808	TTGACTTTCAGGAGTATGAATATGATGATGATGATGATGATGATGATGATGATGATG	1867		
Db	1832	TTGATTTTCAGAGATATGAATATGATGATGATGATGATGATGATGATGATGATGATG	1891		
Qy	1868	TTGGGAGGCTCTGGGCTCTGGGCTCTGGGAGGATGATGATGATGATGATGATGATGATG	1927		
Db	1892	TTGGGAGGCTCTGGGCTCTGGGCTCTGGGAGGATGATGATGATGATGATGATGATGATG	1951		
Qy	1928	TTAGTAAACCGGAGTCTCAATTTCAAGTGGCGTGAAGATGCTTAAAGAGAAAGCTGACA	1987		
Db	1952	TTAGTAAACCGGAGTCTCAATTTCAAGTGGCGTGAAGATGCTTAAAGAGAAAGCTGACA	2011		
Qy	1988	GCTGTGAAAAGAGCTCTCATGTCGAGCTCAAAATGATGATGATGATGATGATGATGATG	2047		
Db	2012	GCTGTGAAAAGAGCTCTCATGTCGAGCTCAAAATGATGATGATGATGATGATGATGATG	2071		
Qy	2048	ACAACATCGTGAATCTGCTGGGCGCATGACACTGTGAGGCGGAGTGTGATTTTGTG	2107		
Db	2072	AGAATATTGTGACCTGTGCGGGCGTGCACACTGTGAGGCGGAGTGTGATTTTGTG	2131		

```
QY 2108 AATATTGTTGCTATGTCACCTCCTCAACTACCTAAGAAAGTAAAGAGAGAAAGTTTCACA 2167
DB 2132 AATACTGTTGCTATGTTGATCTTCTCAACTATCTAAGAGTAAAGAGAAATTTTACA 2191
QY 2168 GGACATGGACAGAGATTTTAAAGGAACATAAATTTCACTTCTTACCTACTTTCCAGGCAC 2227
DB 2192 GGACTTGGACAGAGATTTTCAAGGAACACAAATTTTCACTTTTACCCCACTTTTCAATCAC 2251
QY 2228 ATTCAAAATCCAGCATGCCCTGGTTTCAGAGAGATTCAGTTACACCCGCCCTTGGATCAGC 2287
DB 2252 ATCCAAATTCAGCATGCCCTGGTTTCAAGAGAGATTCAGATACACCCGCACTTCGGATCAAA 2311
QY 2288 TCTCAGGGTTCAATGGGAATTCATCTCTGAAGATGAGATTCATATGAAATGAAACACAGA 2347
DB 2312 TCTCAGGGCTTCATGGGAATTCATCTCAGCTGAAGATGAAATTCATATGAAACACAA 2371
QY 2348 AGAGGCTGGCAGAGAGAGAGAGATTTGAACGCTGTCAGCTTTGAAGACCTCTCTTT 2407
DB 2372 AAAGGCTG-----GAAGAAGAGGAGGACTTGAATGTCTTACATTTGAAGATCTTCTT 2425
QY 2408 GCTTTGGCTACCAAGTGGCCAAAGGCATGGAAATTCCTGGAGTTCAAGTCGTGTGCACA 2467
DB 2426 GCTTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGGATTTAAGTCGTGTGTTCACA 2485
QY 2468 GAGACCTGGCAGCAGGAATGTGTGTGTCACCCAGGAGGTGTGAAGATCTGTGACT 2527
DB 2486 GAGACCTGGCCGCGCAGGAGCTGTGTGTCACCCAGGGAAGTGTGAAGATATGTGACT 2545
QY 2528 TTGGACTGGCCGCGGACATCCTGAGCGACTCCAGCTACGTCGTCAGGGGCAACGACGGC 2587
DB 2546 TTGGATTGGCTCGAGATATCATGATGATTCACACTATGTGTGTCAGGGCAATGCCGTC 2605
QY 2588 TGCGGTTGAATGGATGGCACCAGAGACTTATTTGAAGGGATTCACAAATCAAGAGTG 2647
DB 2606 TGCGTGTAAAATGGATGGCCCCGAAAGCCCTGTTTGAAGGCATCTACACCAATTAAAGTG 2665
QY 2648 ACGTCTGCTCCTACGGCATCTCTCTCTGGGAGATATTTTCACTGGGTGTGAACCTTACC 2707
DB 2666 ATGTCTGGCTCATATGGAATATTAAGTGTGGGAAATCTTCTCACTGGGTGTGAATCCTTACC 2725
QY 2708 CTGGCAATCTCTGTCGACGCTAACTTCTATAAATGATTCAGAGTGGATTTAAATGGAGC 2767
DB 2726 CTGGCAATCTCGGTTGATGCTAACTTCTACAACTGATTCACAAATGGATTTAAATGGATC 2785
QY 2768 AGCAATCTATGCCACAGAGAGGATATACCTTTGTAATGCAATCTCTGCTGGCTTTTGACT 2827
DB 2786 AGCAATTTATGTACAGAGAAATATACATTAATATGCAATCTCTGCTGGCTTTTGACT 2845
QY 2828 CAAGGAAGCGGCCATCTTCCCCAACCTGACTTTCATTTTATAGGATGTACAGTGCAGAGG 2887
DB 2846 CAAGGAAGCGGCCATCTTCCCTAATTTGACTTCGTTTTAGGATGTACAGTGCAGATG 2905
QY 2888 CAGAAGAG-----CATGTATCAGAACATCCATCCATCTACC 2924
DB 2906 CAGAAGAGCGATGTATCAGAAATGTGATGSCCGTGTTCGGAATGTCTTACACACCTACC 2965
QY 2925 AAAACAGGCGGCCCTCAGCAGAGAGGCGGCTCAGAGC-CCAGTGCAGCAGCGCCAGG 2983
DB 2966 AAAACAGGCGACCTTTCAGCAGAGAGATGGATTTGGGGCTACTCTCTCCGCAAGCTCAGG 3025
QY 2984 T-GAAGATTACAGAGAAAGATTTAGCGAGGAGCGCTTGAGCC-----CGCCACCTAG 3038
DB 3026 TCGAAGATTTCGATAGGAACAAATTTAGTTTAAAGGACTTCATCTCCCTACCATATCCCTAA 3085
QY 3039 CAGGCTGTAGACCGACAGCCAGATTAAGCTCTGCCCTCT--GAGGAAGCGCCCTACAGCG 3096
DB 3086 CAGGCTGTAGATTACAAAACAGATTAATTTCACTACATAAAGAAATCTATTATCAAC 3145
QY 3097 CGTTGCTCGCTGACGATTTTCTCTAGATGCTGTCTGCCATTACTC-----CAAGTGA 3149
DB 3146 TGCTGCTTCACGACATTTTCTCTAGAGCGCTCTGCGCTTTTACTCTTTGTTTCAAGGGA 3205
```

```
QY 3150 CTTCTATAAATCAAAACCTCTCTCGCACAGCGGGAGAGCAATAATGAGACTTGTGG 3209
DB 3206 CTTTGTGAAAATCAAAATCATCTGTGCACAAAGCAGGAGAGCTGATAATGAACTTTATG 3265
QY 3210 TGACCCCGCTTACCCTGGGGCTTTCCACAGCTTTGAGGGAAAGCCATGTATCTGAA 3269
DB 3266 GAGCATTGATCTGATCCAAAGCCCTCTCAGCGCGCTTGAGTGAATTTGTACCTTGAAG 3325
QY 3270 TATAGTATATCTTGTAAATACGTGAAACCAACCAACCCGTTTTTTTGTAAAGGAAAGC 3329
DB 3326 TACAGTATATCTTGTAAATACATAAAACAAA-----AGCATTTTGTCTAGGAGAACG 3378
QY 3330 TAAATATGATTTTAAATAATCTATGTTTTTAAATACTATGTAACCTTTTTCATCTATTAG 3389
DB 3379 TAAATGATTTTTT--AAGTCTATCTTTTAAATAATATGTAATTTTTCAGCTATTAG 3436
QY 3390 TGATATATTTTATGGATGGAATAAATCTTCTACTGTAAA 3429
DB 3437 TGATATATTTTATGGTGGGAATAAATTTCTACTACAGA 3476

RESULT 13
LOCUS I44733 3475 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 3 from patent US 5635388.
ACCESSION I44733
VERSION I44733.1 GI:2469446
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3475)
AUTHORS Bennett,B.D., Broz,S.D., Matthews,W. and Zeigler,F.C.
TITLE Agonist antibodies against the f1k2/flt3 receptor and uses thereof
JOURNAL Patent: US 5635388-A 3 03-JUN-1997;
FEATURES
Location/Qualifiers
1..3475
/organism="unknown"
BASE COUNT 1042 a 709 c 784 g 940 t
ORIGIN

Query Match 63.8%; Score 2203; DB 6; Length 3475;
Best Local Similarity 79.6%; Pred. No. 0;
Matches 2749; Conservative 0; Mismatches 650; Indels 56; Gaps 10;

QY 8 GCCTACCGCGCTCCCGAGGCCATCGCGGCGTTGGCGAGCGCAGCGACCGCGGCTGC 67
DB 35 GGGGACCCCGGCTCCCGAGGCCATCGCGGCGTTGGC---GCGGACGCGGCGCACCGTGC 91
QY 68 TGCTGCTTGTGTTTGTTCAGTAAATGATCTTGAGACCGTTTACAAACCAAGACCTGCCTG 127
DB 92 CGCTGCTGTTGTTTCTGCAATGATATTTGGGACTATTTACAATCAAGATCTGCCTG 151
QY 128 TGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGTTCATCAGCGGGAAGCCATCAT 187
DB 152 TGATCAAGTGTGTTTAAATCAATCAATAAGAACAAATGATTCATCAGTGGGGAAGTCATCAT 211
QY 188 CGTACCGAATGGTGGGAGGATCCCGAGAGACCTCCAGTGTACCCGCGGAGCCGAGAGTG 247
DB 212 CATATCCCATGGTATCAGAAATCCCGGAGACCTCGGGTGTGCGTTGAGACCCGAGAGGT 271
QY 248 AAGGACGGTATATGAAGCGCCACCGTGGAGTGGCGGAGTCTGGGTCCATCACCTGC 307
DB 272 CAGGACAGTGTACGAAGCTGCCGCTGGGAAGTGGATGTATCTGCTCCATCACACTGC 331
QY 308 AAGTGCAGCTGCCACCCCGGAGGACCTTCTCGCTCTGGGTCTTTTAAGCAGAGCTCC 367
DB 332 AAGTGCAGTGTGAGGAGTGTCTGCTCTGGGTCTTTTAAGCAGAGCTCC 391
QY 368 TGGGCTGCCAGCGGCACTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCCATCTGA 427
DB 392 TGAATTGCCGACATTTTGTATTTACAAAACAGAGGAGTCTGTTCCATGGTCAATTTGA 451
```









QY	368	TGGGCTGCCAGCGCCACTTTGATTTACAAAACAGAGGAATCGTTCCATGTGGCCATCTTGA	427	QY	1448	CCAATTGCAGCGAGGAATCCCGAAGAGGAGTTGGTAATAAAAGGCTAACAGAAAAGTGT	1507
Db	392	TGAATTGCCAGCCACATTTGATTTACAAAACAGAGGAGTTGTTCCATGTCTATTTGA	451	Db	1472	CCAACCTGCACACAGAGATCAGAGAGGAGTCTGGAATAGAAAGGCTAACAGAAAAGTGT	1531
QY	428	ACGTGACAGAGACCCAGGAGAGAAATACCTACTCCATATTCAGAGCGAAGCGCAACT	487	QY	1508	TTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGAAGCCGGGAAAGGCTTCTGG	1567
Db	452	AAATGACAGAAACCCCAAGCTGGAGAAATACCTACTTTTATTCAGAGTGAAGCTACCAATT	511	Db	1532	TTGGACAGTGGGTGTCGAGCAGTACTCTAAACATGAGTGAAGCCATAAAAGGTTCTCTGG	1591
QY	488	ACACAGTACTGTTACAGTGAATGAAGAGATACACAGCTGTATGTCTAAGGACCTT	547	QY	1568	TCAAAATGCTGCTGCTACAAATTTATGGCAGCGTCTTGGAAACCATCTTTTAAACTCAC	1627
Db	512	ACACAATATGTTTACAGTGAATGAAGAAATACCCCTGCTTTACACATTAAGAAGACCTT	571	Db	1592	TCAGTGTGTCATACAAATCCCTTGGCCACATCTTGTGAGAGCATCTTTTAAACTCTC	1651
QY	548	ACTTTAGGAAGATGAAAACACAGGATGCACCTCTGCTATCCGAGGCTGTTCCGGAGC	607	QY	1628	CAGGCCCTTCCCTTTTCATCCAAAGACACATCTCTTCTATGCGACCATTTGGCTCTGTC	1687
Db	572	ACTTTAGAAAATGAAAACACAGGAGCCCTGCTGCTATCTCAGAGCGTTCCAGAGC	631	Db	1652	CAGGCCCTTCCCTTTTCATCCAAAGACACATCTCTATGCAACAATTTGGTGTGTC	1711
QY	608	CACTGTGGAGTGTCTGTCAGCTCCACAGGAAAGCTGTAAAGAAGAGCCCTG	667	QY	1688	TCCCTTCAATGTTGTTCTCATTTGTTGATCGCCACAAATACAAAAGCAATTTAGGT	1747
Db	632	CGATCTGGAATGGTGTCTTCGATTCACAGGGGAAAGCTGTAAAGAAGAAAGTCCAG	691	Db	1712	TCCTCTTCAATGTTGTTTAAACCCTGCTAATTTGTCAAGTACAAAAGCAATTTAGGT	1771
QY	668	CTGTTGTCAGAAAGAGGAAAAGTACTTTCATGACTTGTTCGGAACACATCAGATGCT	727	QY	1748	ACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCCCTGGATACGAGTACTTCTACG	1807
Db	692	CTGTTGTTAAAGAGGAAAAGTGTCTCATGAATTTTGGGACGACATGAAGTGCT	751	Db	1772	ATGAAGCCAGCTACAGATGGTACAGTGACCGGCTCTTCAGATATGATGACTTCTACG	1831
QY	728	GTCTAGAAAATGCACCTGGCCGCGAATGCACCAAGCTGTTCAACCATAGATCTAAACACAGG	787	QY	1808	TTGACTTCAGGAGCTATGAATATGACCTTAAGTGGGAGTTCCTCCGAGAGAGAACTTAGAGT	1867
Db	752	GTGCGCAAAATGAACTGGCGAGGAATGCACCGAGCTGTTCAATATGATCTTAATCAAA	811	Db	1832	TTGATTTTCAGAGAATATGAATATGATCTCAAAATGGGAGTTTCCAAAGAGAAAATTTAGAGT	1891
QY	788	CTCCTCAGAGCACACTGCCCCAGTTATTCTCGAAAAGTGGGGAACCCCTTGTGGATCAGGT	847	QY	1868	TTGGGAAGTCTCTGGGCTCTGGGCTTTCGGGAGGTGATGAACGCCACGGCTTATGGCA	1927
Db	812	CTCCTCAGACACATTTGCCACAATTTATTTTAAAGTAGGGGAACCCCTTATGGATAAGGT	871	Db	1892	TTGGGAAGGTACTAGGATCAGGTGCTTTTGGAAAAGTGTGAACGCAACAGCTTATGGAA	1951
QY	848	GTAAGCCATCCATGTGAACCATGATTCGGGCTCACCTGGAGAGCTGGAAGCAAGCCC	907	QY	1928	TTAGTAAACAGGAGTCTCAATTCAGGTGCGGTGAAGATGCTTAAAGAGAAAGCTCACA	1987
Db	872	GCAAAAGCTGTTATGTGAACCATGATTCGGGCTCACCTGGGAATTAGAAAACAAAGCAC	931	Db	1952	TTAGCAAAACAGAGTCTCAATCCAGGTTGCGGTCAAAATGCTGGAAGAAAAGAGCAGACA	2011
QY	908	TGGAGAGGGGAGCTACTTTGAGATGAGTACCTACTCCACAACAGGACCATGATTTCGGA	967	QY	1988	GTGTGAAAAGAAAGCTCTCATGTGCGAGCTCAAAATGATGACCCACTGGGACACCATG	2047
Db	932	TCGAGAGGGGCACTCTTTGAGATGAGTACCTATTCAACAACAGAACTATGATACGGA	991	Db	2012	GTCTGAAAGAGAGGCACTCATGTGAGAATCAAGATGATGACCCAGCTGGGAAGCCACG	2071
QY	968	TTCTCTTGGCCCTTTGTCTTCCGTGGGAAGAACGACACCCGGATATTACACCTGCTCTT	1027	QY	2048	ACAACATCTGTAATCTGTGGGGGCAATGCACACTGTCAGGGCCAGTGTACTGATTTTTG	2107
Db	992	TTCTCTTGTCTTTGTATCATCAGTGGCAAGAAACGACACCCGGATACACTTGTCTCT	1051	Db	2072	AGAATATTGTGAACCTGCTGGGGGCGTGACACTGTCAAGAACCAATTTACTTGATTTTTG	2131
QY	1028	CCTCAAGCCACCCAGCCAGTCAGGTTGGTGACCATCTCAGAAAAGGGTTTAAAGC	1087	QY	2108	ATATATTGTGCTATGTTGACCTCCCAACTACTTAAGAGTAAAGAGAGAACTTTTACACA	2167
Db	1052	CTTCAAGCATCCCAAGTCAATCAGCTTTGGTTACCATCGTAGGAAGGGATTTATAAATG	1111	Db	2132	AAATCTGTTGCTATGTTGATCTTCTCAACTATCTAAGAGTAAAGAGAGAAATTTTACACA	2191
QY	1088	CTACAGCTCGCAAGAGATGATAAATGACCCCTAGCAAAAGTTCTGCTCTCAGTCA	1147	QY	2168	GGACATGGACAGAGATTTTAAAGGAACATAAATTCAGTTCTTACCCCTACTTTCCAGGCAC	2227
Db	1112	CTACCAATTCAGTGAAGATTAAGAAATGACCAATATGAAGATTTGTTTTCTGTCA	1171	Db	2192	GGACTTGGACAGAGATTTCAAGGAACACAAATTCAGTTTACCCCACTTTCCCAATCAC	2251
QY	1148	GGTTTAAAGCCGTACCCAGGAATCCGATGCACGTGGATCTTCTCAAGCCCTCATTTCCCTT	1207	QY	2228	ATTCAAATTCAGATGCTGTTTCCAGAGAAAGTTCAGTTTACCCGCCCTTGGATCAGC	2287
Db	1172	GGTTTAAAGCCGTACCCACAAATCAGATGTACGTGGACCTTCTCTCGAAAATCATTTCTCT	1231	Db	2252	ATTCAAATTCAGATGCTGTTTCAAGAGAAAGTTCAGATACACCCGGACTCTCGATCAAA	2311
QY	1208	GTGAACAGAGAGGCTCGAGGATGGGTACAGCATATCTAAATTTGGCATATAGACACA	1267	QY	2288	TCTCAGGGTTCATGGGAATTCAAATTCATCTGAAAGATGAGATTGATGAAAACCCAGA	2347
Db	1232	GTGAGCAAAAGGGTCTTGATACGGATACAGCATATCCAAAGTTTGTGCAATCATAGACACC	1291	Db	2312	TCTCAGGGTTCATGGGAATTCATTTCACTCTGAAGTGAATTTGAATATGAAAACCCAAA	2371
QY	1268	AGCCAGGAGATACATATCTATGAGAAATGATGACCCCGATGACCCCAATGTTCA	1327	QY	2348	ACAGGCTGCGCAGAAAGAGGAGGAAGATTGAACGCTGCTGACGTTTGAAGACCTCTCTT	2407
Db	1292	AGCCAGGAGATATATATTTCCATGAGAAAATGATGATGCCCAATTTACCAAAATGTTCA	1351	Db	2372	AAAGCGTG-----GAAGAGAGGAGGACTTGAATGTGCTTACATTTGAAGATCTCTTTT	2425
QY	1328	CGCTGAATATAGAAGAAACCTCAAGTGTAGCAAAATGCCCTCAGCCAGCCAGCGTCT	1387	QY	2408	GCTTTTCCGTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCTGTGTCCACA	2467
Db	1352	CGCTGAATATAGAAGAAACCTCAAGTGTCTCGAGAGCATCGCAGAGTCAAGGCTCT	1411	Db	2426	GCTTTGATATCAAGTGGCCAAAGGAATGGAATTTCTGGAAATTTAAGTCTGTGTTCACA	2485
QY	1388	GTTCCTCTGATGGCTACCCGCTTACCTCTTGGACCTGGAAGAAGTGTTCGGCAAAATCTC	1447	QY	2468	GAGACCTGCGACGCCAGGAATGTTGGTCAACCCACGGGAGGTTGAGATCTGTGACT	2527
Db	1412	GTTCCTCGGATGATACCATTTACCATCTTGGACCTTGGAGAAAGTGTTCAGACAAGTCTC	1471	Db	2486	GAGACCTGCGCCGAGGAACGCTGTGTCAACCCACGGGAAAGTGTGAGATATGACT	2545
				QY	2528	TTGGACTGCGCCCGAGACATCTTCAGCGACTTCAGCTACGTCAGCGGGCAACGCCAGGC	2587





Search completed: May 25, 2003, 09:18:59  
Job time : 8670.83 secs

Db 2337 CTTGAATGCTGTACATTGAGATCTCTCTTTGCTTTGCATATCAAGTTGCCAAAGGAAT 2396  
Qy 2436 GGAATTCCTGGAGTTCAAGTGTGTCCACAGACACCTGGCAGCCAGGAATGTCTGGT 2495  
Db 2397 GGAATTTCTGAAATTAAGTGTGTGTCCACAGACACCTGGCCCGCAGAACGCTGCTGT 2456  
Qy 2496 CACCCACGGGAAGTGGTGAAGATCTGTGACTTTGGACTGGCCCGCAGACATCCTGAGCGA 2555  
Db 2457 CACCCACGGGAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTGA 2516  
Qy 2556 CTCAGCTACGTCGACGGGCAAGCCAGCGCTGCCGTTGAAGTGGATGGCACCCGAGAG 2615  
Db 2517 TTCCAACTATGTGTGACGGGCAATGCCCGTCTGCTGTAAATGGATGGCCCGCAAG 2576  
Qy 2616 CTTATTTGAAGGATCTACAAATCAAGAGTGACGTCTGGCTACGGCATCCTTCTCTG 2675  
Db 2577 CTTGTTGAAGGCATCTACACCATTAAGAGTGTCTGCGTCATATGGAATATTACTGTG 2636  
Qy 2676 GGAGATATTTTCACTGGGTGTGAACCCCTTACCCTGGCATTCCTGTGACGCTAACTTCTA 2735  
Db 2637 GGAATCTTCTCACTTGTGTGATCCTTACCCTGGCATTCGGTGTGATGCTAACTTCTA 2696  
Qy 2736 TAACTGATTCAGAGTGAATTAATAAGGAGCAGCCATTCATGCCACAGAGGATATA 2795  
Db 2697 CAAACTGATTCAAATGATTTAAATGGATCAGCCATTTTATGCTACAGAAGAAATATA 2756  
Qy 2796 CTTTGTATGCAATCCTGCTGGCTTTTGAAGAGGAGCGGCATCCTTCCCAACCT 2855  
Db 2757 CATTATATGCAATCCTGCTGGCTTTTGAAGAGGAGCGGCATCCTTCCCAATTT 2816  
Qy 2856 GACTTCATTTTTAGGATGTCAGCTGGCAGAGGCGCAGAAG- 2896  
Db 2817 GACTTCGTTTTAGGATGTCAGCTGGCAGATGCAGAGAAGCGATGTATCAGATGTGGA 2876  
Qy 2897 ----CATGTATCAGAACATCCATCCATCTACAAACAGCGGCCCTCAGCAGAGAGGC 2952  
Db 2877 TGGCCGTGTTTCGGAATGTCTCACACCTACCAAAACAGCGGACCTTTCAGCAGAGAT 2936  
Qy 2953 GGGCTCAGAGC-CCAGTCGCCACAGCGCCAGGT-GAAGATTCACAGAGAAAGATTAGC 3010  
Db 2937 GGATTTGGGGTACTCTCTCCGACGGCTCAGGTCGAGATTCGTAGAGGAACAATTTAGT 2996  
Qy 3011 GAGGAGCCTTGGACCC---CGCCACCTAGCAGGCTGTAGACCGCAGAGCCAAAGATTA 3066  
Db 2997 TTTAAGGACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTA 3056  
Qy 3067 GCCTCGCCTCT--GAGGAAGCGCCCTACAGCGGCTTGCCTCGTGGACTTTTCTCTAGAT 3124  
Db 3057 ATTTTCATCACTAAAGAAATCTATTATCAACTGCTGCTTCACCAGACTTTTCTCTAGAA 3116  
Qy 3125 GCTGCTGCCATTACTC-----CAAAGTCACCTTCTATAAATCAARACC-TCTCCTCGC 3176  
Db 3117 GCTGCTGCGTTTACTCTGTTTTCAAAGGACATTTGTAAATCAAAATCATCCTGTAC 3176  
Qy 3177 ACAGCGGAGAGCAATATGAGACTGTGTTGTGAGCCGCTTACCCTGGGGGCTTTC 3236  
Db 3177 AAGCAGGAGAGCTGATATGAACTTTATTGGAGCATGTATCTGCATCCAAGGCTTCT 3236  
Qy 3237 CACGAGCTTGAGGGGAAGCCATGATCTGAAATATAGTATATTCTTTGTAATAACGTGAA 3296  
Db 3237 CAGGCTGGCTTGAGTGAATG-TGTACCTGAAGTACAGTATATTCTTTGTAATAACATAAA 3295  
Qy 3297 ACAACCAACCCCTTTTTCCTAAGGGAAGCTAAATATGATTTTAAAAATCTATGTT 3356  
Db 3296 ACAA-----AGCATTTTGTCTAAGGAGAGCTAAATGATTTTTT--AAGCTATGTT 3346  
Qy 3357 TTAATACTATGTAATTTTTTTCATCTATTAGTGATATATTTTATGATGGAATAAAC 3416  
Db 3347 TTAATAATATGTAATAATTTTTCAGCTATTAGTGATATATTTTATGGGGGAATAAAA 3406  
Qy 3417 TTTCTACTG 3425  
Db 3407 TTCTACCG 3415

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: May 25, 2003, 04:06:12 : Search time 695.168 Seconds  
(without alignments)  
11185.997 Million cell updates/sec

Title: US-09-919-408-1  
Perfect score: 3453  
Sequence: 1 GCGGCTGGTACGCGGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :				N_Geneseq_101002:*			
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*						
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*						
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*						
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*						
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*						
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*						
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*						
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*						
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*						
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*						
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*						
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*						
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*						
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*						
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*						
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*						
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*						
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*						
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*						
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*						
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*						
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*						
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*						
24:	/SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*						

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3453	100.0	3453	16	AAQ81012
2	3453	100.0	3453	17	AAQ81012
3	3453	100.0	3453	18	AAQ81012
4	3453	100.0	3453	20	AAQ81012
5	3451.4	100.0	3453	14	AAQ81012
6	3451.4	100.0	3453	14	AAQ81012
7	3451.4	100.0	3453	16	AAQ81012
8	3448.2	99.9	3453	13	AAQ81012
9	3446.6	99.8	3453	14	AAQ81012

10	3342.4	96.8	3521	16	AAQ80801	Flk2/fit3 tyrosine
11	2227.2	64.5	3501	16	AAQ81013	Flk2 receptor prot
12	2227.2	64.5	3501	16	AAQ79069	Human flk-2 cDNA.
13	2227.2	64.5	3501	17	AAQ79069	Murine foetal live
14	2227.2	64.5	3501	18	AAQ72117	Human flk-2 recep
15	2227.2	64.5	3501	20	AAQ77515	Human flk-2 cDNA.
16	2225.6	64.5	3501	14	AAQ53503	Human flk-2 cDNA.
17	2225.6	64.5	3501	14	AAQ40915	Human flk-2 cDNA.
18	2224	64.4	3501	14	AAQ35250	Human flk-2 coding
19	2204.8	63.9	3476	16	AAQ91536	Human STK-1 cDNA.
20	2198.2	63.7	3475	16	AAQ08002	Human Flk2/fit3 ty
21	2085	60.4	3120	14	AAQ49756	PTK gene LPTK25.
22	2081.8	60.3	3120	16	AAQ03096	Protein tyrosine-k
23	2024.6	58.6	2949	19	AAQ39041	Human receptor typ
24	2018.8	58.5	2958	19	AAQ39042	Human receptor typ
25	2017.2	58.4	2958	19	AAQ39040	Human receptor typ
26	1994.8	57.8	2982	19	AAQ39039	Human receptor typ
27	1984.2	57.5	2978	19	AAQ39038	Human receptor typ
28	1199.6	34.7	1894	15	AAQ54036	Flk-2ws gene. Mus
29	1119.6	32.4	2247	19	AAQ55294	Sequence PHON32390
30	357.4	10.4	3992	23	AAQ79666	DNA encoding novel
31	357.4	10.4	3992	24	AAQ48105	Human macrophage c
32	332.6	9.6	3069	21	AAQ24718	Bovine c-Kit bk-1
33	331.4	9.6	2919	20	AAQ80687	Hampshire porcine
34	329	9.5	5084	19	AAQ20443	Human c-kit Oncoge
35	329	9.5	5084	22	AAQ13426	Human Kit/stem cel
36	329	9.5	5084	24	ABL64113	Breast cancer rela
37	329	9.5	5084	24	ABL64113	Ovary cancer relat
38	325.4	9.4	6390	23	AAQ79665	DNA encoding novel
39	325.4	9.4	6390	23	AAQ84936	DNA encoding novel
40	325	9.4	5098	22	AAQ13425	Murine Kit/stem ce
41	286.4	8.3	2463	20	AAQ89336	Platelet-derived g
42	286.4	8.3	4054	11	AAQ06869	Sequence encodes p
43	286.4	8.3	4544	10	AAQ90355	cDNA encoding plat
44	286.4	8.3	4544	17	AAQ34552	Platelet-derived g
45	286.4	8.3	5602	23	AAQ84940	DNA encoding novel

ALIGNMENTS

RESULT 1				AAQ81012 standard; cDNA; 3453 BP.			
ID	AAQ81012	standard	cDNA	3453	BP.		
XX	AAQ81012						
AC	AAQ81012						
DT	16-AUG-1995	(first entry)					
DE	Flk2 receptor protein-tyrosine-kinase cDNA.						
XX	Mouse Flk2; receptor protein-tyrosine-kinase; primitive						
KW	hematopoietic cell; fetal liver kinase; ds.						
XX	Mus musculus.						
OS							
XX	Key	Location/Qualifiers					
FT	CDS	31..3009					
FT		/*tag= a					
FT		/product= Flk2 receptor protein-tyrosine-kinase					
FT	sig_peptide	31..111					
FT		/tag= b					
FT	mat_peptide	112..3006					
FT		/*tag= c					
XX	WO9500554-A.						
XX	05-JAN-1995.						
XX	17-JUN-1994;	94WO-US06944.					
XX	18-JUN-1993;	93US-0080244.					
XX	21-JUN-1993;	93US-0081508.					

PR 23-NOV-1993; 93US-0157490.  
XX (UYPR-) UNIV PRINCETON.  
XX Lemischka IR;  
XX WPI: 1995-052014/07.  
DR P-PSDB; AAR67815.  
XX  
PT Ligand for receptor protein tyrosine kinase - useful for the  
PT stimulation of primitive haematopoietic stem cells causing  
PT proliferation and/or differentiation  
XX  
XX Disclosure; Fig 1a; 131pp; English.  
XX  
CC The sequence corresponds to a cDNA encoding a mouse Flk2 (fetal  
CC liver kinase) receptor protein-tyrosine-kinase. Flk2 is expressed  
CC in primitive hematopoietic cells but not in mature hematopoietic  
CC cells. The gene product is useful in isolation of receptor ligands,  
CC which have applications in diagnosis of bone marrow disorders and in  
CC stimulating proliferation and/or differentiation of primitive  
CC hematopoietic stem cells.  
XX  
SQ Sequence 3453 BP; 947 A; 822 C; 875 G; 809 T; 0 other;  
Query Match 100.0%; Score 3453; DB 16; Length 3453;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCGGCTGTGCTACCGCGCTCCGAGGCGCATCGGGCGTTGGCGAGCGACGCGCG 60  
DB 1 GCGGCTGTGCTACCGCGCTCCGAGGCGCATCGGGCGTTGGCGAGCGACGCGCG 60  
QY 61 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
DB 61 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
QY 121 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
DB 121 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
QY 181 CCATCATCTACCGAATGCTGCGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGCGC 240  
DB 181 CCATCATCTACCGAATGCTGCGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGCGC 240  
QY 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCCATC 300  
DB 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCCATC 300  
QY 301 ACCCTGCAAGTGCAGCTGCCACCCAGGGGACCTTTCCTGCTGCTGCTGCTGCTGCT 360  
DB 301 ACCCTGCAAGTGCAGCTGCCACCCAGGGGACCTTTCCTGCTGCTGCTGCTGCTGCT 360  
QY 361 AGCTCCCTGGGCTGCCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 361 AGCTCCCTGGGCTGCCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
QY 421 ATCTTGAACGTGACAGACCCAGGACGAGGAGATACCTACTCCATATTCAGACGCGC 480  
DB 421 ATCTTGAACGTGACAGACCCAGGACGAGGAGATACCTACTCCATATTCAGACGCGC 480  
QY 481 GCCAATACACACTACTGTTACAGTGAATGTAAGAGATACACAGTGTATGCTTAAGG 540  
DB 481 GCCAATACACACTACTGTTACAGTGAATGTAAGAGATACACAGTGTATGCTTAAGG 540  
QY 541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCTGCTGCTGCTGCTGCTGCTGCT 600  
DB 541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCTGCTGCTGCTGCTGCTGCTGCT 600  
QY 601 CCGGAGCCCACTGTGGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
DB 601 CCGGAGCCCACTGTGGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

QY 661 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
DB 661 GGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
QY 721 AGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
DB 721 AGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
QY 781 AACCAGGCTCCTCAGAGCAGCAGTCCCGCAGTATTCCTGAAAGTGGGGACCTTCTGG 840  
DB 781 AACCAGGCTCCTCAGAGCAGCAGTCCCGCAGTATTCCTGAAAGTGGGGACCTTCTGG 840  
QY 841 ATCAGGTGTAAGGCCATCCATGTAACCATGATGCGGCTCAGCTGGAGCTGGAAGAC 900  
DB 841 ATCAGGTGTAAGGCCATCCATGTAACCATGATGCGGCTCAGCTGGAGCTGGAAGAC 900  
QY 901 AAGCCCTCGAGAGGCGAGCTACTTTGAGATGAGTACCTACTCCCAACAGGACCATG 960  
DB 901 AAGCCCTCGAGAGGCGAGCTACTTTGAGATGAGTACCTACTCCCAACAGGACCATG 960  
QY 961 ATTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
DB 961 ATTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
QY 1021 TGCTCTTCTCAAAAGCACCCAGCCAGCTCAGCGTTGGTGACCATCTCTAGAAAAGGTTT 1080  
DB 1021 TGCTCTTCTCAAAAGCACCCAGCCAGCTCAGCGTTGGTGACCATCTCTAGAAAAGGTTT 1080  
QY 1081 ATAAAGCTACAGCTGCGCAAGAGAGTATGAATTCACCGTACGAAAAGTCTTCTGCTTC 1140  
DB 1081 ATAAAGCTACAGCTGCGCAAGAGAGTATGAATTCACCGTACGAAAAGTCTTCTGCTTC 1140  
QY 1141 TCAGTCAGTGTAAAGGTTACCCAGCAATCCGATGCGACGCTGGATCTCTCTCAAGCTCA 1200  
DB 1141 TCAGTCAGTGTAAAGGTTACCCAGCAATCCGATGCGACGCTGGATCTCTCTCAAGCTCA 1200  
QY 1201 TTTCTCTGTGAACAGAGAGCGCTGGAGATGGGTACAGCATATCTAAATTTTGGCGATCAT 1260  
DB 1201 TTTCTCTGTGAACAGAGAGCGCTGGAGATGGGTACAGCATATCTAAATTTTGGCGATCAT 1260  
QY 1261 AAGAACAAGCCAGGAGAGTACATATTCTATGCGAGAAATGATGAGCGCCAGTTCACCAAA 1320  
DB 1261 AAGAACAAGCCAGGAGAGTACATATTCTATGCGAGAAATGATGAGCGCCAGTTCACCAAA 1320  
QY 1321 ATCTTCAGCTGATATGAAGAGACCTCAAGTGTAGCAATGCTCAGCAGCGCCAG 1380  
DB 1321 ATCTTCAGCTGATATGAAGAGACCTCAAGTGTAGCAATGCTCAGCAGCGCCAG 1380  
QY 1381 GCGTCTGTTCTCTGATGGCTACCCGCTACCCCTCTTGGAGCTGGAAGAGTGTTCGGAC 1440  
DB 1381 GCGTCTGTTCTCTGATGGCTACCCGCTACCCCTCTTGGAGCTGGAAGAGTGTTCGGAC 1440  
QY 1441 AATCTCCCAATTCAGCGAGGAAATCCAGAGAGGTTTGGAAATAAAGGCTAACAGA 1500  
DB 1441 AATCTCCCAATTCAGCGAGGAAATCCAGAGAGGTTTGGAAATAAAGGCTAACAGA 1500  
QY 1501 AAGTGTGTTGGCAGTGGGTGTCAGCAGTACTCTAAATATGAGTGGCGCGGGAAGGG 1560  
DB 1501 AAGTGTGTTGGCAGTGGGTGTCAGCAGTACTCTAAATATGAGTGGCGCGGGAAGGG 1560  
QY 1561 CTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
DB 1561 CTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620  
QY 1621 AACTCAGCAGGCGCTTCCCTTTTATCCAAAGACATCTCTCTATGCGACCATTTGGG 1680  
DB 1621 AACTCAGCAGGCGCTTCCCTTTTATCCAAAGACATCTCTCTATGCGACCATTTGGG 1680  
QY 1681 CTCTGTCTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
DB 1681 CTCTGTCTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
QY 1741 TTTAGGTACGAGAGTGCAGTATCCAGGTGACTGCGCCCTTGGGCTGCTGCTGCTGCTGCTGCT 1800

```
Db 1741 TTTAGTACGAGAGTCAAGTGCAGATGATCAGGTGACTGCGCCCTCGATAACGAGTAC 1800
Qy 1801 TTCTACCTTGACTTCAGGAGTATGAATATGACCTTAAGTGGAGTTCCTCCAGAGAGAAC 1860
Db 1801 TTCTACCTTGACTTCAGGAGTATGAATATGACCTTAAGTGGAGTTCCTCCAGAGAGAAC 1860
Qy 1861 TTAGAGTTTGGAAAGTCCCTGGGGTCTGGGGCTTTTCGGGAGGGTGATCAACGCCACGCGCC 1920
Db 1861 TTAGAGTTTGGAAAGTCCCTGGGGTCTGGGGCTTTTCGGGAGGGTGATCAACGCCACGCGCC 1920
Qy 1921 TATGGCATTAGTAAACGGGAGTCTCAATTGAGTGGCGGTCAAGATGCTATAAGAGAGAAA 1980
Db 1921 TATGGCATTAGTAAACGGGAGTCTCAATTGAGTGGCGGTCAAGATGCTATAAGAGAGAAA 1980
Qy 1981 GCTCAGAGCTGTGAAGAAGAGTCTCATGTCGAGCTCAAAATGATGACCCAGCTGGGA 2040
Db 1981 GCTCAGAGCTGTGAAGAAGAGTCTCATGTCGAGCTCAAAATGATGACCCAGCTGGGA 2040
Qy 2041 CACCATGACAACTCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCCAGTGTACTTG 2100
Db 2041 CACCATGACAACTCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCCAGTGTACTTG 2100
Qy 2101 ATTTTGAATATTGTTGATGTTGACCTCTCAACTACCTAAGAAACATAATTTTCACTTCCCTACTTTC 2160
Db 2101 ATTTTGAATATTGTTGATGTTGACCTCTCAACTACCTAAGAAACATAATTTTCACTTCCCTACTTTC 2160
Qy 2161 TTTTACAGGACATGGACAGAGATTTTAAAGAAACATAATTTTCACTTCCCTACTTTC 2220
Db 2161 TTTTACAGGACATGGACAGAGATTTTAAAGAAACATAATTTTCACTTCCCTACTTTC 2220
Qy 2221 CAGGCATTCACAAATCCAGCATGCTGGTTCAGGAGAGTTCAGTTACACCCGCCCTTG 2280
Db 2221 CAGGCATTCACAAATCCAGCATGCTGGTTCAGGAGAGTTCAGTTACACCCGCCCTTG 2280
Qy 2281 GATCAGCTCTCAGGGTTCAATGGGAATTCATTTCTGAAGATGAGATTCAATATGAA 2340
Db 2281 GATCAGCTCTCAGGGTTCAATGGGAATTCATTTCTGAAGATGAGATTCAATATGAA 2340
Qy 2341 AACCAGAGAGCTGGCAGAGAGAGAGAGAGAGATTTGACAGTCTGACGTTTGAAGAC 2400
Db 2341 AACCAGAGAGCTGGCAGAGAGAGAGAGAGAGATTTGACAGTCTGACGTTTGAAGAC 2400
Qy 2401 CTCCTTTGCTTGGTACCAAGTGGCCAAAGCATGGAATTCCTGGAGTTCGAAGTCGCTG 2460
Db 2401 CTCCTTTGCTTGGTACCAAGTGGCCAAAGCATGGAATTCCTGGAGTTCGAAGTCGCTG 2460
Qy 2461 GTCCACAGAGACCTGGCAGCAGGAATGTGTTGGTCAACCCAGGGAAGTGGTGAAGATC 2520
Db 2461 GTCCACAGAGACCTGGCAGCAGGAATGTGTTGGTCAACCCAGGGAAGTGGTGAAGATC 2520
Qy 2521 TGTGACTTTGACTTGGCCCGAGACATCTGAGCAGTCCAGCTAGCTGCTCAGGGGCAAC 2580
Db 2521 TGTGACTTTGACTTGGCCCGAGACATCTGAGCAGTCCAGCTAGCTGCTCAGGGGCAAC 2580
Qy 2581 GCAGGCTGCGGTGAAGTGTGATGSCACCCAGAGCTTATTTGAGGATCTACACAATC 2640
Db 2581 GCAGGCTGCGGTGAAGTGTGATGSCACCCAGAGCTTATTTGAGGATCTACACAATC 2640
Qy 2641 AAGAGTACGCTGCTGCTACGGCATCCTTCTCTGGGAGATATTTTCACTGGGTGTAAC 2700
Db 2641 AAGAGTACGCTGCTGCTACGGCATCCTTCTCTGGGAGATATTTTCACTGGGTGTAAC 2700
Qy 2701 CCTTACCTTGGCATCTGTCGAGCTAACTTCTATAAACTGATTCAGAGTGGATTAAA 2760
Db 2701 CCTTACCTTGGCATCTGTCGAGCTAACTTCTATAAACTGATTCAGAGTGGATTAAA 2760
Qy 2761 ATGGAGAGCCATCTATGCCACAGAGGGATATCTTTGTAATCCAATCTGCTGGGCT 2820
Db 2761 ATGGAGAGCCATCTATGCCACAGAGGGATATCTTTGTAATCCAATCTGCTGGGCT 2820
Qy 2821 TTTGACTCAAGGAAGCGGCCATCTTCCCAACCTGACTTCAATTTTAGGATGTCAGCTG 2880
```

```
Db 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTCAATTTTAGGATGTCAGCTG 2880
Qy 2881 CGAGGGCAGAGAGAGCATGTATCAGAACATCCATCCATCTACCAAAACAGGGGCCCT 2940
Db 2881 CGAGGGCAGAGAGAGCATGTATCAGAACATCCATCCATCTACCAAAACAGGGGCCCT 2940
Qy 2941 CAGCAGAGAGGGCGCTCAGAGCCAGTGCACACAGCGCCAGGTGAAGATTACAGAGAA 3000
Db 2941 CAGCAGAGAGGGCGCTCAGAGCCAGTGCACACAGCGCCAGGTGAAGATTACAGAGAA 3000
Qy 3001 AGAAGTTAGCGAGAGGCTTGGAGCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCCA 3060
Db 3001 AGAAGTTAGCGAGAGGCTTGGAGCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCCA 3060
Qy 3061 AGATTAGCTCGCCTCTGAGGAAGCGCCTACAGCGCGTTCCTCGCTGGACTTTTCTCT 3120
Db 3061 AGATTAGCTCGCCTCTGAGGAAGCGCCTACAGCGCGTTCCTCGCTGGACTTTTCTCT 3120
Qy 3121 AGATGCTGTGCTCCATTACTCCAAAGTGAATCTTATAAAATCAAACTCTCTCGCACAG 3180
Db 3121 AGATGCTGTGCTCCATTACTCCAAAGTGAATCTTATAAAATCAAACTCTCTCGCACAG 3180
Qy 3181 GCGGAGAGCCCAATATGAGACTTGTGTGAGCCCGCTACCTGGGGCCTTTCCACG 3240
Db 3181 GCGGAGAGCCCAATATGAGACTTGTGTGAGCCCGCTACCTGGGGCCTTTCCACG 3240
Qy 3241 AGCTTGAGGGGAAGCCATGTATCTGAAATAGTATATTTTGAATAACGTGAACAA 3300
Db 3241 AGCTTGAGGGGAAGCCATGTATCTGAAATAGTATATTTTGAATAACGTGAACAA 3300
Qy 3301 ACCAAACCCGTTTGTGTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTTAA 3360
Db 3301 ACCAAACCCGTTTGTGTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTTAA 3360
Qy 3361 AATACTATGTAATCTTTTCACTATTTAGTATATATTTATGATGGAATAAACTTTC 3420
Db 3361 AATACTATGTAATCTTTTCACTATTTAGTATATATTTATGATGGAATAAACTTTC 3420
Qy 3421 TACTGTAAAAAAGGAAAAAAAAAAAAAAAAAAAA 3453
Db 3421 TACTGTAAAAAAGGAAAAAAAAAAAAAAAAAAAA 3453
```

## RESULT 2

AAT38733

ID AAT38733 standard: cDNA: 3453 BP.

XX

AC AAT38733;

XX

DT 11-DEC-1996 (first entry)

XX

DE Human foetal liver kinase 2 cDNA.

XX

Human; foetal liver kinase 2; flk-2; protein tyrosine kinase;  
monoclonal; antibody; extracellular domain; receptor assay;  
haematopoietic stem cell; ligand; stimulation; proliferation;  
differentiation; treatment; anaemia; bone marrow damage;  
cancer chemotherapy; radiation; ds.

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS

FT sig\_peptide

FT mat\_peptide

FT US5548065-A.

XX

XX 20-AUG-1996.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX





Db	1561	CTTCTGTCACAAATGCTGTGCGTCAAAATTCATGGCCACGCTTTCGGAAGAACCATCTTTT	1620
Qy	1621	AACTCACAGGCCCCCTTCCCTTTTCATCCAAAGACAACATCTCCTCTCTATGCGACCATGGG	1680
Db	1621	AACTCACAGGCCCCCTTCCCTTTTCATCCAAAGACAACATCTCCTCTCTATGCGACCATGGG	1680
Qy	1681	CTCTGCTCCCTTCATGTGTTCTCATTTGTTGTTCTCATTTGTTGTTGATCTGCGCCAAATACAAAAGCAA	1740
Db	1681	CTCTGCTCCCTTCATGTGTTCTCATTTGTTGTTCTCATTTGTTGTTGATCTGCGCCAAATACAAAAGCAA	1740
Qy	1741	TTTAGGTACGAGAGTCAGCTCGACATGATCCAGTGACTGCGCCCTCGGATACAGGATAC	1800
Db	1741	TTTAGGTACGAGAGTCAGCTCGACATGATCCAGTGACTGCGCCCTCGGATACAGGATAC	1800
Qy	1801	TTCTACGTTTGACTTTCAGGACTATGAATATGACCTTAAGTGGAGTTCCCGACAGAGAAC	1860
Db	1801	TTCTACGTTTGACTTTCAGGACTATGAATATGACCTTAAGTGGAGTTCCCGACAGAGAAC	1860
Qy	1861	TTAGAGTTTGGGAAGGTCCTGGGGCTCGGGCTTTTCGGGAGGTGATGAACGCCACGGCC	1920
Db	1861	TTAGAGTTTGGGAAGGTCCTGGGGCTCGGGCTTTTCGGGAGGTGATGAACGCCACGGCC	1920
Qy	1921	TATGGCATTAGTAAACGGGAGTCTCAATTCAGTGGCGGTGAAGATGCTTAAAGAGAAA	1980
Db	1921	TATGGCATTAGTAAACGGGAGTCTCAATTCAGTGGCGGTGAAGATGCTTAAAGAGAAA	1980
Qy	1981	GCTGCACAGCTGTGAAAAAGAGCTCTCATGCTCGGAGCTCAAAATGATGACCCACCTGGGA	2040
Db	1981	GCTGCACAGCTGTGAAAAAGAGCTCTCATGCTCGGAGCTCAAAATGATGACCCACCTGGGA	2040
Qy	2041	CACCATGACAACTGCGAATCTGCTGGGGCATGCACACTGTGAGGGCCAGTGTACTTG	2100
Db	2041	CACCATGACAACTGCGAATCTGCTGGGGCATGCACACTGTGAGGGCCAGTGTACTTG	2100
Qy	2101	ATTTTTCATATTTGCTATGTGACTCCTCAACTACCTAGAAGTAAAGAGAGAAG	2160
Db	2101	ATTTTTCATATTTGCTATGTGACTCCTCAACTACCTAGAAGTAAAGAGAGAAG	2160
Qy	2161	TTTTCACAGGACATGGACAGAGATTTTAAAGGAACATAATTTCAAGTCTTTACCCCTACTTTC	2220
Db	2161	TTTTCACAGGACATGGACAGAGATTTTAAAGGAACATAATTTCAAGTCTTTACCCCTACTTTC	2220
Qy	2221	CAGGCACATTCAAATTCACAGCATGCTGTTTCACGAGAAGTTCAGTTCACCCGCCCTTG	2280
Db	2221	CAGGCACATTCAAATTCACAGCATGCTGTTTCACGAGAAGTTCAGTTCACCCGCCCTTG	2280
Qy	2281	GATCAGCTCTCAGGGTTCATGGGAATTCAAATTCATCTGAAGATGAGATGGAATGAA	2340
Db	2281	GATCAGCTCTCAGGGTTCATGGGAATTCAAATTCATCTGAAGATGAGATGGAATGAA	2340
Qy	2341	AACCAAGAAGGCTGGCAGAAAGAGAGGAAGATTTGAACGTGCTGACGTTTGAAGAC	2400
Db	2341	AACCAAGAAGGCTGGCAGAAAGAGAGGAAGATTTGAACGTGCTGACGTTTGAAGAC	2400
Qy	2401	CTCCTTTGCTTTGCGTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCTG	2460
Db	2401	CTCCTTTGCTTTGCGTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCTG	2460
Qy	2461	GTCCACAGAGACCTGGCAGCCAGGAATGTTTGGTCAACCCACGGGAAGGTGGTGAAGATC	2520
Db	2461	GTCCACAGAGACCTGGCAGCCAGGAATGTTTGGTCAACCCACGGGAAGGTGGTGAAGATC	2520
Qy	2521	TGTGACTTTGGACTGGCCCGGAGACATCCTCAGCGACTTCCAGCTACGTCGTCAGGGCAAC	2580
Db	2521	TGTGACTTTGGACTGGCCCGGAGACATCCTCAGCGACTTCCAGCTACGTCGTCAGGGCAAC	2580
Qy	2581	GCACGGCTGCCGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTACACAATC	2640
Db	2581	GCACGGCTGCCGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGATCTACACAATC	2640
Qy	2641	AAGAGTGACGCTGTGGTCTTACGGCATCCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700

D	b	2641	AAGAGTGACGCTCTGGTCCCTACGGCATCCCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700
Q	y	2701	CCTTACCCTGGCATTCCTCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGGATTAAA	2760
D	b	2701	CTTTACCCTGGCATTCCTCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGGATTAAA	2760
Q	y	2761	ATGGAGCAGCCATTCTATGCCCACAGAGGGATATCTTTGTAATGCAATCCTCTGGGCT	2820
D	b	2761	ATGGAGCAGCCATTCTATGCCCACAGAGGGATATCTTTGTAATGCAATCCTCTGGGCT	2820
Q	y	2821	TTTGACTACGGAAGCGGCCATCTCTCCCAACCTGACTTCATTTTAGGATGTACGCTG	2880
D	b	2821	TTTGACTACGGAAGCGGCCATCTCTCCCAACCTGACTTCATTTTAGGATGTACGCTG	2880
Q	y	2881	GCAGAGCCAGAAGCATGTATCAGAACATCCATCCATCTACCAAAACAGGGGGCCCT	2940
D	b	2881	GCAGAGCCAGAAGCATGTATCAGAACATCCATCCATCTACCAAAACAGGGGGCCCT	2940
Q	y	2941	CAGCAGAGAGCGGGCTCAGAGCCCAGTCCGCACAGCGCCAGGTGAAGATTACACAGAA	3000
D	b	2941	CAGCAGAGAGCGGGCTCAGAGCCCAGTCCGCACAGCGCCAGGTGAAGATTACACAGAA	3000
Q	y	3001	AGAAGTTAGCGAGGAGCCCTTGGACCCGCCACCCCTAGCAGGCTGTAGACCGCAGAGCCA	3060
D	b	3001	AGAAGTTAGCGAGGAGCCCTTGGACCCGCCACCCCTAGCAGGCTGTAGACCGCAGAGCCA	3060
Q	y	3061	AGATTAGCCCTCGCTCTGAGGAAGCGCCCTACAGCGCTTGTCTCGCTGGACTTTTCTCT	3120
D	b	3061	AGATTAGCCCTCGCTCTGAGGAAGCGCCCTACAGCGCTTGTCTCGCTGGACTTTTCTCT	3120
Q	y	3121	AGATGCTGTCTGCCATTACTTCCAAAGTGACTTCTATAAAATCAAACTCTCTCTCGCACAG	3180
D	b	3121	AGATGCTGTCTGCCATTACTTCCAAAGTGACTTCTATAAAATCAAACTCTCTCTCGCACAG	3180
Q	y	3181	CGGGAGAGCCAATATGAGACTTGTGTGTGAGCCCGCCCTACCCCTGGGGCCCTTCCACG	3240
D	b	3181	CGGGAGAGCCAATATGAGACTTGTGTGTGAGCCCGCCCTACCCCTGGGGCCCTTCCACG	3240
Q	y	3241	AGCTTGAGGGGAAGCCATGTATCTGAAATATAGTATATTTCTTGTAAATACGTGAAACAA	3300
D	b	3241	AGCTTGAGGGGAAGCCATGTATCTGAAATATAGTATATTTCTTGTAAATACGTGAAACAA	3300
Q	y	3301	ACCAAAACCCGTTTTTTGCTAAGGGAAGCTAAATATGATTTTTTAAAAATCTATGTTTTAA	3360
D	b	3301	ACCAAAACCCGTTTTTTGCTAAGGGAAGCTAAATATGATTTTTTAAAAATCTATGTTTTAA	3360
Q	y	3361	AATACTATGTAACTTTTTTCATCTATTTAGTGATATTTTTTATGGATGGAATAAATCTTC	3420
D	b	3361	AATACTATGTAACTTTTTTCATCTATTTAGTGATATTTTTTATGGATGGAATAAATCTTC	3420
Q	y	3421	TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA	3453
D	b	3421	TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA	3453

DEC 11 1963

RESULTS  
AD72118

AA172118  
ID AAT72118 standard: cDNA: 3453 BP.

OT 7 / 1111 22

AC AAT72118;

XX  
XX

- 3 -

DT 19-AUG-1997 (first entry)

XX  
-----

DE Murine flk-2 receptor coding sequence.

XX

KW Human; fetal liver kinase 2; flk2; receptor protein tyrosine kinase;

KW pTK; liver; spleen; thymus; adult; brain; bone marrow; primitive portion;

KW haematopoietic

KW proliferation

KW macrocytic an

XX

OS Mus musculus.

XX

FH	Key	Location/Qualifiers	
CDS		31..3009	
FT		/*tag= a	
FT		/product= Murine_flk-2	
FT	sig_peptide	31..111	
FT		/*tag= b	
FT	mat_peptide	112..3006	
FT		/*tag= c	
XX			
XX	US5621090-A.		
PN			
XX			
XX			
PD	15-APR-1997.		
XX			
PF	02-APR-1991;	91US-0679666.	
XX			
PR	26-JUN-1992;	92US-0906397.	
PR	02-APR-1991;	91US-0679666.	
PR	28-JUN-1991;	91US-0728913.	
PR	15-NOV-1991;	91US-0793065.	
PR	24-DEC-1991;	91US-0813593.	
XX			
XX	(UYPR-) UNIV PRINCETON.		
PI			
PI	Lemischka IR;		
XX			
XX			
DR	WPI; 1997-235228/21.		
DR	P-PSDB; AAW19874.		
XX			
PPT	Protein containing the extracellular domain of human flk-2 - used		
PPT	for identification of primitive haematopoietic cell proliferation		
PPT	and differentiation stimulatory ligands, e.g. for treating anaemia		
XX			
XX	Disclosure; Fig 1A; 55pp; English.		
XX			
CCC	This sequence encodes the murine fetal liver kinase 2 (flk2). flk-2 is		
CCC	a receptor protein tyrosine kinase (pTK) and is important in transducing		
CCC	putative self-renewal signals from the environment. flk-2 is expressed		
CCC	in fetal liver, spleen, fetal thymus, adult brain and adult bone marrow,		
CCC	and it is thought that flk-2 is expressed in the entire primitive portion		
CCC	of the haematopoietic hierarchy. The invention concerns a recombinant		
CCC	nucleic acid, preferably mRNA, which encodes a protein containing only		
CCC	the extracellular domain of human flk-2 and lacking the flk-2 intra-		
CCC	cellular catalytic domain. The resultant protein represents a soluble		
CCC	form of flk-2 which is used to isolate specific ligands for flk-2. These		
CCC	ligands can be used to stimulate proliferation and/or differentiation of		
CCC	mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for		
CCC	treatment of macrocytic or aplastic anaemia or bone marrow damage caused		
CCC	by cancer treatment or radiation.		
XX			
XX	Sequence 3453 BP; 947 A; 822 C; 875 G; 809 T; 0 other;		
SQ			
	Query Match	100.0%; Score 3453; DB 18; Length 3453;	
	Best Local Similarity	100.0%; Pred. No. 0;	
	Matches 3453; Conservative	0; Mismatches	0; Gaps
		0; Indels	0;
QY	1	GGCGCCTGGCTACCGCGCTCCGGAGGCCATCGGCGGTGGCGCAGCAGCACCGG	60
Db	1	GGCGCCTGGCTACCGCGCTCCGGAGGCCATCGGCGGTGGCGCAGCAGCACCGG	60
QY	61	CGGCTGCTGCTGCTGTTGTTTGTCTAGTAATGATTCTTGAGACCGTTACAAACCAAG	120
Db	61	CGGCTGCTGCTGCTGTTGTTTGTCTAGTAATGATTCTTGAGACCGTTACAAACCAAG	120
QY	121	CTGCGCTGATCAGTGTGTTTAAATCAGTCATCAGAACAAATGCTCATCAGCGGGAAG	180
Db	121	CTGCGCTGATCAGTGTGTTTAAATCAGTCATCAGAACAAATGCTCATCAGCGGGAAG	180
QY	181	CCATCATCGTACCGAATGGTCGAGGATCCCGAAGACCTCCAGTGATCCCGCAGCGCG	240
Db	181	CCATCATCGTACCGAATGGTCGAGGATCCCGAAGACCTCCAGTGATCCCGCAGCGCG	240
QY	241	CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGTGCCCGAGTCTGGGTCCATC	300
Db			

Qy	1381	CGGTCTCTGTTCTCTGATGGCTACCCCGCTACCCCTCTTGAGACCTTGGAGAAAGTGTTCGGAC	1440
Db	1381	GCSTCTCTGTTCTCTGATGTCTACCCGCTACCCCTCTTGACCTTGGAGAAAGTGTTCGGAC	1440
Qy	1441	AAATCTCCCAATTCGACGGAGGAATCCCAAGAGGAGTTTGGAAATAAAGGCGTAACAGA	1500
Db	1441	AAATCTCCCAATTCGACGGAGGAATCCCAAGAGGAGTTTGGAAATAAAGGCGTAACAGA	1500
Qy	1501	AAAGTGTTCGCCAGTGGGTGTCGACGAGTACTCTAAATATGAGTCAGGCGCGGAAGGG	1560
Db	1501	AAAGTGTTCGCCAGTGGGTGTCGACGAGTACTCTAAATATGAGTCAGGCGCGGAAGGG	1560
Qy	1561	CTTCTGGTCAAAATGCTGTGCGTACAATTCATGGGCACGCTCTTCGGAACCATCTTTTAA	1620
Db	1561	CTTCTGGTCAAAATGCTGTGCGTACAATTCATGGGCACGCTCTTCGGAACCATCTTTTAA	1620
Qy	1621	AACTCACAGGCGCCCTTCCTTCATCCAAGACAACATCTCCTTCATGTCGACCATTTGGG	1680
Db	1621	AACTCACAGGCGCCCTTCCTTCATCCAAGACAACATCTCCTTCATGTCGACCATTTGGG	1680
Qy	1681	CTCTGTCTCCCTTCATTTGTTCTCATTTGTTCATTTGTTGATCTGCCACAATAACAAAAGCAA	1740
Db	1681	CTCTGTCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAATAACAAAAGCAA	1740
Qy	1741	TTTAGGTACAGAGTCAGCTGCAGATGATCCAGTGACTGGCCCTCTGGATACGAGTAC	1800
Db	1741	TTTAGGTACAGAGTCAGCTGCAGATGATCCAGTGACTGGCCCTCTGGATACGAGTAC	1800
Qy	1801	TTCTACGTTGCATTCAGGGACTATGAATATGACCTTAAAGTGGAGTTCCTCGAGAGAGAAC	1860
Db	1801	TTCTACGTTGCATTCAGGGACTATGAATATGACCTTAAAGTGGAGTTCCTCGAGAGAGAAC	1860
Qy	1861	TTAGAGTTTGGGAAGTCTCTGGGTCTGGCGCTTTGGGAGGTTGATGAACGCCACGGCC	1920
Db	1861	TTAGAGTTTGGGAAGTCTCTGGGTCTGGCGCTTTGGGAGGTTGATGAACGCCACGGCC	1920
Qy	1921	TATGGCAATTAGTAAACGGGAGTCTCAATTCAGTGGCGTGAAGATGCTTAAAGAGAAA	1980
Db	1921	TATGGCAATTAGTAAACGGGAGTCTCAATTCAGTGGCGTGAAGATGCTTAAAGAGAAA	1980
Qy	1981	GCTGCACGCTGTGAAAAAAGACTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA	2040
Db	1981	GCTGCACGCTGTGAAAAAAGACTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA	2040
Qy	2041	CACATGACAACTGCGTAATCTGCTGGGGCATGCACACTGTACAGGGCCAGTGTACTTG	2100
Db	2041	CACATGACAACTGCGTAATCTGCTGGGGCATGCACACTGTACAGGGCCAGTGTACTTG	2100
Qy	2101	ATTTTTCGAATATGTTCTATGTTGACCTCCTCAACTACCTTAAGAGTAAAGAGAGAAG	2160
Db	2101	ATTTTTCGAATATGTTCTATGTTGACCTCCTCAACTACCTTAAGAGTAAAGAGAGAAG	2160
Qy	2161	TTTTCACAGGACATGGACAGAGATTTTAAAGSAACATAATTTACGTTCTTACCCTACTTTC	2220
Db	2161	TTTTCACAGGACATGGACAGAGATTTTAAAGSAACATAATTTACGTTCTTACCCTACTTTC	2220
Qy	2221	CAGGCACATTCAAATTCACGATGCCTGTTTACGAGAGAAGTTCAGTTACACCGCCCTTG	2280
Db	2221	CAGGCACATTCAAATTCACGATGCCTGTTTACGAGAGAAGTTCAGTTACACCGCCCTTG	2280
Qy	2281	GATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATTCGAGATGAGATTGAATGAA	2340
Db	2281	GATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATTCGAGATGAGATTGAATGAA	2340
Qy	2341	AACCAGAAGAGGCTGGCAGAGAAGAGGAGGAAGATTGAAACGTGCTGACGTTTGAAGAC	2400
Db	2341	AACCAGAAGAGGCTGGCAGAGAAGAGGAGGAAGATTGAAACGTGCTGACGTTTGAAGAC	2400
Qy	2401	CTCCTTTGCTTTGGGTACCAAGTGGCGCAAGGCAATGGAATTCCTGGAGTTCAAGTCGT	2460
Db	2401	CTCCTTTGCTTTGGGTACCAAGTGGCGCAAGGCAATGGAATTCCTGGAGTTCAAGTCGT	2460

RESULT 4  
AAX77514  
ID AAX  
XX

AC AAX77514;  
 XX DT  
 XX DE Murine flk-2 cDNA.  
 XX DE  
 XX Murine; flk-2; cell isolation; fetal liver kinase; receptor;  
 KW monoclinal; polyclonal; antibody; tyrosine kinase; ds.  
 XX OS Mus sp.  
 XX FT Key  
 FT CDS 31..3009  
 FT Location/Qualifiers  
 FT /\*tag= a  
 FT /product= "flk-2"  
 XX US5912133-A.  
 XX 15-JUN-1999.  
 XX 10-FEB-1998; 98US-0021324.  
 XX 19-NOV-1992; 92US-0977451.  
 XX 02-APR-1991; 91US-0679666.  
 XX 28-JUN-1991; 91US-0728913.  
 XX 15-NOV-1991; 91US-0793065.  
 XX 24-DEC-1991; 91US-0813593.  
 XX 26-JUN-1992; 92US-0906397.  
 XX 12-NOV-1992; 92US-0975049.  
 XX 30-APR-1993; 93US-0055269.  
 XX 31-OCT-1994; 94US-0252498.  
 XX 15-FEB-1996; 96US-0601891.  
 XX (UYPR-) UNIV PRINCETON.  
 XX Lemischka IR;  
 XX WPI; 1999-357194/30.  
 XX P-PSDB; AAY08616.  
 XX Isolating hematopoietic cells expressing fetal liver kinase 1  
 PT receptors  
 XX Disclosure; Fig 1a; 59pp; English.  
 XX This invention describes a novel method of isolating cells expressing  
 CC fetal liver kinase 1 (flk-1) receptors on their surface and comprises  
 CC binding the cells to a polyclonal or monoclonal antibody specific to  
 CC the FLK-1 receptor and isolating the cells that have bound to the  
 CC antibody. The method can be used to isolate hematopoietic stem cells in  
 CC any mammal but preferably a rat, mouse, rabbit or human. The proteins of  
 CC the invention belong to the receptor protein family. This sequence  
 CC encodes the murine flk-2 protein which is used in the method of the  
 CC invention.  
 XX SQ Sequence 3453 BP; 947 A; 822 C; 875 G; 809 T; 0 other;  
 Query Match 100.0%; Score 3453; DB 20; Length 3453;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCGGCTGGCTACCGCGCTCCGGAGGCATCGGGCGTTGGCGAGCGCAGCGACCGG 60  
 DB 1 GCGGCTGGCTACCGCGCTCCGGAGGCATCGGGCGTTGGCGAGCGCAGCGACCGG 60  
 QY 61 CGGCTGCTGCTGTTGTTGTTGTCAGTAATGATCTTGAGACCGGTTACAAACCAAGAC 120  
 DB 61 CGGCTGCTGCTGTTGTTGTTGTCAGTAATGATCTTGAGACCGGTTACAAACCAAGAC 120  
 QY 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGAAAG 180  
 DB 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGAAAG 180

181 CCATCATCTACCGAATGGTGGAGGATCCCGAAGACCTCCAGTGTACCCCGAGGCGC 240  
 DB  
 181 CCATCATCTACCGAATGGTGGAGGATCCCGAAGACCTCCAGTGTACCCCGAGGCGC 240  
 QY 241 CAGAGTGAAGGACGGGTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCTGGGTCCATC 300  
 DB 241 CAGAGTGAAGGACGGGTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCTGGGTCCATC 300  
 QY 301 ACCTCGCAAGTGCAGCTCGCCACCCAGAGGACCTTTCTGCTCTGGTCTTTAAAGCAC 360  
 DB 301 ACCTCGCAAGTGCAGCTCGCCACCCAGAGGACCTTTCTGCTCTGGTCTTTAAAGCAC 360  
 QY 361 AGCTCCCTGGGCTGCCAGCGCCACTTTGATTACAAAACAGAGAAATCGTTTCCATGGCC 420  
 DB 361 AGCTCCCTGGGCTGCCAGCGCCACTTTGATTACAAAACAGAGAAATCGTTTCCATGGCC 420  
 QY 421 ATCTTGAACGTGACAGAGACCCAGGAGGAGAAATACCTACTTCATATTCAGAGGGAACGC 480  
 DB 421 ATCTTGAACGTGACAGAGACCCAGGAGGAGAAATACCTACTTCATATTCAGAGGGAACGC 480  
 QY 481 GCCAACTACACAGTACTGTTACAGTCAATCTAGAGATACACAGCTGTATGCTAAGC 540  
 DB 481 GCCAACTACACAGTACTGTTACAGTCAATCTAGAGATACACAGCTGTATGCTAAGC 540  
 QY 541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCCTGCTGCTGCTATCTCCGAGGGTGT 600  
 DB 541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCCTGCTGCTGCTATCTCCGAGGGTGT 600  
 QY 601 CCGGAGCCCACTGTGGAGTGGTGTCTGTCAGCTCCACAGGAAAGCTGTAAAGAGAA 660  
 DB 601 CCGGAGCCCACTGTGGAGTGGTGTCTGTCAGCTCCACAGGAAAGCTGTAAAGAGAA 660  
 QY 661 GGCCCTGCTGTTGTCAGAAAGGAGAAAGTACTTCTATGATGTTGTCGGAACAGACATC 720  
 DB 661 GGCCCTGCTGTTGTCAGAAAGGAGAAAGTACTTCTATGATGTTGTCGGAACAGACATC 720  
 QY 721 AGATGCTGTCTAGAAATGCACCTGGGCGCGGAATSCACCAAGCTGTTTCCACCATAGATCTA 780  
 DB 721 AGATGCTGTCTAGAAATGCACCTGGGCGCGGAATSCACCAAGCTGTTTCCACCATAGATCTA 780  
 QY 781 AACAGGCTCCTCAGAGCACACTGCCCCAGTATTCTGAAAGTGGGGGAAACCTTTGTGG 840  
 DB 781 AACAGGCTCCTCAGAGCACACTGCCCCAGTATTCTGAAAGTGGGGGAAACCTTTGTGG 840  
 QY 841 ATCAGGTGAAGGCCATCCATGTGAACCATGGATTTCGGGCTCACTGGAGCTGGAACAC 900  
 DB 841 ATCAGGTGAAGGCCATCCATGTGAACCATGGATTTCGGGCTCACTGGAGCTGGAACAC 900  
 QY 901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 960  
 DB 901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 960  
 QY 961 ATTCCGATCTCTTGGCCTTTGTCCTTCGTTGGGAAGGAACAGACACCGGATATTACACC 1020  
 DB 961 ATTCCGATCTCTTGGCCTTTGTCCTTCGTTGGGAAGGAACAGACACCGGATATTACACC 1020  
 QY 1021 TGCTCTCTCCAAAGCACCCAGCCAGTGCAGCTGGTGGTGCATCTCTAGAAAAGGGTGT 1080  
 DB 1021 TGCTCTCTCCAAAGCACCCAGCCAGTGCAGCTGGTGGTGCATCTCTAGAAAAGGGTGT 1080  
 QY 1081 ATAAACCTACCAGCTCGCAAGAGAGATGATAAATTGACCGTACGAAAAGTCTGCTTC 1140  
 DB 1081 ATAAACCTACCAGCTCGCAAGAGAGATGATAAATTGACCGTACGAAAAGTCTGCTTC 1140  
 QY 1141 TCAGTCAAGTGTAAAGCGTACCCACGAATCCGATGCGAGTGTCTTCTCAAGSCCTCA 1200  
 DB 1141 TCAGTCAAGTGTAAAGCGTACCCACGAATCCGATGCGAGTGTCTTCTCAAGSCCTCA 1200  
 QY 1201 TTTCCCTGTGAACAGAGAGGCGCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT 1260  
 DB 1201 TTTCCCTGTGAACAGAGAGGCGCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT 1260  
 QY 1261 AAGAACACGCGAGGAGAGTACATATTCTATGCAGAAAATGATGAGCCAGTTCACCAAA 1320

Db	1361	 AAGAACAGCCAGGAGAGTACATATCTATGCGAAATATGACGCCCGAGTTCCACAAA	1320
Qy	1321	ATGTTACGCTGAATATAAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG	1380
Db	1321	ATGTTACGCTGAATATAAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG	1380
Qy	1381	CGTCTCTGTTCTCTGTATGGCTACCCGCTACCCCTCTTGGACCTTGGGAAGTGTTCGGAC	1440
Db	1381	CGTCTCTGTTCTCTGTATGGCTACCCGCTACCCCTCTTGGACCTTGGGAAGTGTTCGGAC	1440
Qy	1441	AAATCTCCCAATGTCACGGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAGGCTAACAGA	1500
Db	1441	AAATCTCCCAATGTCACGGAGGAAATCCCAAGAGGAGTTTGGAAATAAAAGGCTAACAGA	1500
Qy	1501	AAAGTGTTTGGCCAGTGGGTGTCCGACGAGTACTCTAAATATGAGTGGGCGCGGGAAGGG	1560
Db	1501	AAAGTGTTTGGCCAGTGGGTGTCCGACGAGTACTCTAAATATGAGTGGGCGCGGGAAGGG	1560
Qy	1561	CTTCTGGTCAAAATGCTGTGGTACAAATCTATGGGACGCTCTTGGCAAAACCATCTTTTA	1620
Db	1561	CTTCTGGTCAAAATGCTGTGGTACAAATCTATGGGACGCTCTTGGCAAAACCATCTTTTA	1620
Qy	1621	AACTCACCAGGCCCCCTTCCTTCATCCAGACAAATCTCTCTATGCGACCATTTGGG	1680
Db	1621	AACTCACCAGGCCCCCTTCCTTCATCCAGACAAATCTCTCTATGCGACCATTTGGG	1680
Qy	1681	CTCTGTCTCCCTTCATTGTGTCTCATTTGTGATCTGCCACAAATACAAAAAGCAA	1740
Db	1681	CTCTGTCTCCCTTCATTGTGTCTCATTTGTGATCTGCCACAAATACAAAAAGCAA	1740
Qy	1741	TTTAGGTACGAGAGTCAGCTCGAGATGATCCAGGTGACTGGCCCCCTGGATAACGAGTAC	1800
Db	1741	TTTAGGTACGAGAGTCAGCTCGAGATGATCCAGGTGACTGGCCCCCTGGATAACGAGTAC	1800
Qy	1801	TTCTACGTTGACITTCAGGGACTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC	1860
Db	1801	TTCTACGTTGACITTCAGGGACTATGAATATGACCTTAAAGTGGAGTTCCCGAGAGAGAAC	1860
Qy	1861	TTAGAGTTTGGGAAGGTCCTGGGCTCGCGCTTTCGGGAGGTTGATGAACGCCACGGCC	1920
Db	1861	TTAGAGTTTGGGAAGGTCCTGGGCTCGCGCTTTCGGGAGGTTGATGAACGCCACGGCC	1920
Qy	1921	TATGGCATTAGTAAAACGGGAGTCTCAATTACAGTGGCGTGAAGATGCTTAAAGAGAAA	1980
Db	1921	TATGGCATTAGTAAAACGGGAGTCTCAATTACAGTGGCGTGAAGATGCTTAAAGAGAAA	1980
Qy	1981	GCTGACAGCTGTGAAAAGAAAGTCTCATGTCTGGAGCTCAAAATGATGACCCACTGGGA	2040
Db	1981	GCTGACAGCTGTGAAAAGAAAGTCTCATGTCTGGAGCTCAAAATGATGACCCACTGGGA	2040
Qy	2041	CACCATCACAAATCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCCAGTGTACTTG	2100
Db	2041	CACCATCACAAATCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCCAGTGTACTTG	2100
Qy	2101	ATTTTGAATATTGTTGCTATGTTGACCTCTCAACTACCTAAAGAGTAAAAGAGAGAAG	2160
Db	2101	ATTTTGAATATTGTTGCTATGTTGACCTCTCAACTACCTAAAGAGTAAAAGAGAGAAG	2160
Qy	2161	TTTCACAGGACATGGACAGAGATTTTAAAGAACATAATTTCAAGTTCTTACCCTACTTTC	2220
Db	2161	TTTCACAGGACATGGACAGAGATTTTAAAGAACATAATTTCAAGTTCTTACCCTACTTTC	2220
Qy	2221	CAGGCACATTCAAATTCACGATGCCCTGGTTTCAGGAGAAGTTCAGTTTACACCCGCCCTTG	2280
Db	2221	CAGGCACATTCAAATTCACGATGCCCTGGTTTCAGGAGAAGTTCAGTTTACACCCGCCCTTG	2280
Qy	2281	GATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATTCAGATGAGATTGAATATGAA	2340
Db	2281	GATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATTCAGATGAGATTGAATATGAA	2340
Qy	2341	AACCAGAGAGGCTGCGAGAAAGAGGAGGAAGATTGAACGTGCTTGACGTTTGAAGAC	2400
Db	2341	AACCAGAGAGGCTGCGAGAAAGAGGAGGAAGATTGAACGTGCTTGACGTTTGAAGAC	2400

2341	Db	AACACAGAGAGCGCTGGCAGAAAGAGAGGAGGAAGATTTGAACGTGCTGACGCTTTTGAACAC	2401	Qy	CTCCTTTGCTTTTGGCTACCAAGTAGTGGCCAAAGGCATGGAATTTCCCTGGAGTTTCAAGTCGTGT
2401	Db	CTCCTTTGCTTTTGGCTACCAAGTAGTGGCCAAAGGCATGGAATTTCCCTGGAGTTTCAAGTCGTGT	2401	Qy	CTCCTTTGCTTTTGGCTACCAAGTAGTGGCCAAAGGCATGGAATTTCCCTGGAGTTTCAAGTCGTGT
2461	Qy	GTCCACAGAGACCTTGGCAGCCAGAGAAATGTGTTGGTCAACCCACGCGGAAGGTGGTGAAGATC	2520	Qy	GTCCACAGAGACCTTGGCAGCCAGAGAAATGTGTTGGTCAACCCACGCGGAAGGTGGTGAAGATC
2461	Db	GTCCACAGAGACCTTGGCAGCCAGAGAAATGTGTTGGTCAACCCACGCGGAAGGTGGTGAAGATC	2520	Qy	GTCCACAGAGACCTTGGCAGCCAGAGAAATGTGTTGGTCAACCCACGCGGAAGGTGGTGAAGATC
2521	Qy	TGTGACATTTTGACATGGCCCGGAGACATCCTGAGCGACCTCCAGCTACGTCTGTCAGGGGCAAC	2580	Qy	TGTGACATTTTGACATGGCCCGGAGACATCCTGAGCGACCTCCAGCTACGTCTGTCAGGGGCAAC
2521	Db	TGTGACATTTTGACATGGCCCGGAGACATCCTGAGCGACCTCCAGCTACGTCTGTCAGGGGCAAC	2580	Qy	TGTGACATTTTGACATGGCCCGGAGACATCCTGAGCGACCTCCAGCTACGTCTGTCAGGGGCAAC
2581	Qy	GCAGGCTGCCGGTGAAGTGGATGGCACCAGAGAGCTATTTTGAAGGCATCTACACAATC	2640	Qy	GCAGGCTGCCGGTGAAGTGGATGGCACCAGAGAGCTATTTTGAAGGCATCTACACAATC
2581	Db	GCAGGCTGCCGGTGAAGTGGATGGCACCAGAGAGCTATTTTGAAGGCATCTACACAATC	2640	Qy	GCAGGCTGCCGGTGAAGTGGATGGCACCAGAGAGCTATTTTGAAGGCATCTACACAATC
2641	Qy	AAGAGTGACGTCTGGTCTCCTACGGCATCCTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700	Qy	AAGAGTGACGTCTGGTCTCCTACGGCATCCTCTCTGGGAGATATTTTCACTGGGTGTGAAC
2641	Db	AAGAGTGACGTCTGGTCTCCTACGGCATCCTCTCTGGGAGATATTTTCACTGGGTGTGAAC	2700	Qy	AAGAGTGACGTCTGGTCTCCTACGGCATCCTCTCTGGGAGATATTTTCACTGGGTGTGAAC
2701	Qy	CCTTACCGTGGCATTCCTGTCGACGCTAACCTTCTATAAAGTGAATTCAGAGTGGATTTAAA	2760	Qy	CCTTACCGTGGCATTCCTGTCGACGCTAACCTTCTATAAAGTGAATTCAGAGTGGATTTAAA
2701	Db	CCTTACCGTGGCATTCCTGTCGACGCTAACCTTCTATAAAGTGAATTCAGAGTGGATTTAAA	2760	Qy	CCTTACCGTGGCATTCCTGTCGACGCTAACCTTCTATAAAGTGAATTCAGAGTGGATTTAAA
2761	Qy	ATGGAGACGCATTCATGCCCACAGAGAGGATATACTTTGAATGCAATCTCGCTGGGCT	2820	Qy	ATGGAGACGCATTCATGCCCACAGAGAGGATATACTTTGAATGCAATCTCGCTGGGCT
2761	Db	ATGGAGACGCATTCATGCCCACAGAGAGGATATACTTTGAATGCAATCTCGCTGGGCT	2820	Qy	ATGGAGACGCATTCATGCCCACAGAGAGGATATACTTTGAATGCAATCTCGCTGGGCT
2821	Qy	TTTGACTCAAGGAAAGCGGCCATCCTTCCCCAACCTGACTTCATTTTAGGATGTGAGCTG	2880	Qy	TTTGACTCAAGGAAAGCGGCCATCCTTCCCCAACCTGACTTCATTTTAGGATGTGAGCTG
2821	Db	TTTGACTCAAGGAAAGCGGCCATCCTTCCCCAACCTGACTTCATTTTAGGATGTGAGCTG	2880	Qy	TTTGACTCAAGGAAAGCGGCCATCCTTCCCCAACCTGACTTCATTTTAGGATGTGAGCTG
2881	Qy	GCAGAGCAGAGAGCAATGTATCAGAACATCCATCCATCTACCAAAACAGCGCGGCCCT	2940	Qy	GCAGAGCAGAGAGCAATGTATCAGAACATCCATCCATCTACCAAAACAGCGCGGCCCT
2881	Db	GCAGAGCAGAGAGCAATGTATCAGAACATCCATCCATCTACCAAAACAGCGCGGCCCT	2940	Qy	GCAGAGCAGAGAGCAATGTATCAGAACATCCATCCATCTACCAAAACAGCGCGGCCCT
2941	Qy	CAGCAGAGAGCGGGCTCAGAGCCAGTCGCGCCACAGCGCCAGGTGAAGATTACACAGAA	3000	Qy	CAGCAGAGAGCGGGCTCAGAGCCAGTCGCGCCACAGCGCCAGGTGAAGATTACACAGAA
2941	Db	CAGCAGAGAGCGGGCTCAGAGCCAGTCGCGCCACAGCGCCAGGTGAAGATTACACAGAA	3000	Qy	CAGCAGAGAGCGGGCTCAGAGCCAGTCGCGCCACAGCGCCAGGTGAAGATTACACAGAA
3001	Qy	AGAAAGTTAGCAGGAGGCGCTTGGACCCCGCCACCCCTAGCAGGCTGTAGACCGCGAGGCCA	3060	Qy	AGAAAGTTAGCAGGAGGCGCTTGGACCCCGCCACCCCTAGCAGGCTGTAGACCGCGAGGCCA
3001	Db	AGAAAGTTAGCAGGAGGCGCTTGGACCCCGCCACCCCTAGCAGGCTGTAGACCGCGAGGCCA	3060	Qy	AGAAAGTTAGCAGGAGGCGCTTGGACCCCGCCACCCCTAGCAGGCTGTAGACCGCGAGGCCA
3061	Qy	AGATTAGCCTCGCCTCTGAGGAAGCGCCCTACAGCGGCTGCTTCGCTGGAGACTTTTCTCT	3120	Qy	AGATTAGCCTCGCCTCTGAGGAAGCGCCCTACAGCGGCTGCTTCGCTGGAGACTTTTCTCT
3061	Db	AGATTAGCCTCGCCTCTGAGGAAGCGCCCTACAGCGGCTGCTTCGCTGGAGACTTTTCTCT	3120	Qy	AGATTAGCCTCGCCTCTGAGGAAGCGCCCTACAGCGGCTGCTTCGCTGGAGACTTTTCTCT
3121	Qy	AGATGCTGCTGCCATTACTCCAAAGTGACATCTATAAATCAAACTCTCCTCGCCACAG	3180	Qy	AGATGCTGCTGCCATTACTCCAAAGTGACATCTATAAATCAAACTCTCCTCGCCACAG
3121	Db	AGATGCTGCTGCCATTACTCCAAAGTGACATCTATAAATCAAACTCTCCTCGCCACAG	3180	Qy	AGATGCTGCTGCCATTACTCCAAAGTGACATCTATAAATCAAACTCTCCTCGCCACAG
3181	Qy	CGGGAGAGCCAAATAGTAGACTTTTGGTGAGCCCGCCCTACCCTGGGGGCTTTTCCACG	3240	Qy	CGGGAGAGCCAAATAGTAGACTTTTGGTGAGCCCGCCCTACCCTGGGGGCTTTTCCACG
3181	Db	CGGGAGAGCCAAATAGTAGACTTTTGGTGAGCCCGCCCTACCCTGGGGGCTTTTCCACG	3240	Qy	CGGGAGAGCCAAATAGTAGACTTTTGGTGAGCCCGCCCTACCCTGGGGGCTTTTCCACG
3241	Qy	AGCTTGAGGGAAAGCCATGTATCTGAATATATATATTTCTTGTAAATACGTGAACAA	3300	Qy	AGCTTGAGGGAAAGCCATGTATCTGAATATATATATTTCTTGTAAATACGTGAACAA
3241	Db	AGCTTGAGGGAAAGCCATGTATCTGAATATATATATTTCTTGTAAATACGTGAACAA	3300	Qy	AGCTTGAGGGAAAGCCATGTATCTGAATATATATATTTCTTGTAAATACGTGAACAA
3301	Qy	ACCAACCCGTTTTTGGTGAAGGAAAGCTAAATATGATTTTTTAAATCTPATGTTTTTAA	3360	Qy	ACCAACCCGTTTTTGGTGAAGGAAAGCTAAATATGATTTTTTAAATCTPATGTTTTTAA
3301	Db	ACCAACCCGTTTTTGGTGAAGGAAAGCTAAATATGATTTTTTAAATCTPATGTTTTTAA	3360	Qy	ACCAACCCGTTTTTGGTGAAGGAAAGCTAAATATGATTTTTTAAATCTPATGTTTTTAA
3361	Qy	AATACTATGTAACTTTTTCATCTATTTAGTGATATTTTTATGGATGGAAATAAACTTTC	3420	Qy	AATACTATGTAACTTTTTCATCTATTTAGTGATATTTTTATGGATGGAAATAAACTTTC
3361	Db	AATACTATGTAACTTTTTCATCTATTTAGTGATATTTTTATGGATGGAAATAAACTTTC	3420	Qy	AATACTATGTAACTTTTTCATCTATTTAGTGATATTTTTATGGATGGAAATAAACTTTC
3421	Qy	TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453	3421	Qy	TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453
3421	Db	TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453	3421	Qy	TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453





QY	1201	TTTCCTTTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT	1260
DB	1201	TTTTCTTTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT	1260
QY	1261	AAGAACAGCCAGAGAGTACATATTTCTATGCAGAAATGATGACGCCAGTTTCACCAAA	1320
DB	1261	AAGAACAGCCAGAGAGTACATATTTCTATGCAGAAATGATGACGCCAGTTTCACCAAA	1320
QY	1321	ATGTTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG	1380
DB	1321	ATGTTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG	1380
QY	1381	CGCTCTGTTTCTCTGATGGTACCCGCTACCCCTTGGACCTGCGAAGTGTTCGGAC	1440
DB	1381	CGCTCTGTTTCTCTGATGGTACCCGCTACCCCTTGGACCTGCGAAGTGTTCGGAC	1440
QY	1441	AAATCTCCCAATTCACGAGGAAATCCCAAGAGGTTTGAATAAAGGCTAACAGA	1500
DB	1441	AAATCTGCCAATTCACGAGGAAATCCCAAGAGGTTTGAATAAAGGCTAACAGA	1500
QY	1501	AAAGTGTTCGCCAGTGGTCTGAGCAGTACTCTAAATATAGTGAAGCCGGGAAAGGG	1560
DB	1501	AAAGTGTTCGCCAGTGGTCTGAGCAGTACTCTAAATATAGTGAAGCCGGGAAAGGG	1560
QY	1561	CTTCTGTCAAATGCTGTGCTACAAATTTCTATGGCAGCTCTGCGAAACCATCTTTTA	1620
DB	1561	CTTCTGTCAAATGCTGTGCTACAAATTTCTATGGCAGCTCTGCGAAACCATCTTTTA	1620
QY	1621	AACTCACCAGCCCTTCCCTTTTCATCCAGACACATCTCTCTATGCGACCATTTGG	1680
DB	1621	AACTCACCAGCCCTTCCCTTTTCATCCAGACACATCTCTCTATGCGACCATTTGG	1680
QY	1681	CTCTGCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA	1740
DB	1681	CTCTGCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA	1740
QY	1741	TTTAGTACGAGATCAGCTCAGATGATCAGTGAATGAGTGGCCCTCGATTAACAGATAC	1800
DB	1741	TTTAGTACGAGATCAGCTCAGATGATCAGTGAATGAGTGGCCCTCGATTAACAGATAC	1800
QY	1801	TTCTACGTTGACTTCAGGAGTATGAATATGACCTTAAGTGGGAGTTCCCGAGAGAAC	1860
DB	1801	TTCTACGTTGACTTCAGGAGTATGAATATGACCTTAAGTGGGAGTTCCCGAGAGAAC	1860
QY	1861	TTAGAGTTTGGGAAGTCTGGGCTCTGGGCTTTTGGGAGGGTGAACGCCACGGCC	1920
DB	1861	TTAGAGTTTGGGAAGTCTGGGCTCTGGGCTTTTGGGAGGGTGAACGCCACGGCC	1920
QY	1921	TATGGCATTAGTAAACGGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAAAGAAA	1980
DB	1921	TATGGCATTAGTAAACGGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAAAGAAA	1980
QY	1981	GCTGACAGCTGTGAAGAAGCTCTCATGTCGAGCTCAAAATGATGACCCCTGGGA	2040
DB	1981	GCTGACAGCTGTGAAGAAGCTCTCATGTCGAGCTCAAAATGATGACCCCTGGGA	2040
QY	2041	CACCATGACACATCTGTAATCTGCTGGGCGATGACACCTGTCAGGCGCAGTGTACTTG	2100
DB	2041	CACCATGACACATCTGTAATCTGCTGGGCGATGACACCTGTCAGGCGCAGTGTACTTG	2100
QY	2101	ATTTTGAATATTTGCTATGTTGACCTCTCAACTACCTTAAGAAATGAAAGAGAGAG	2160
DB	2101	ATTTTGAATATTTGCTATGTTGACCTCTCAACTACCTTAAGAAATGAAAGAGAGAG	2160
QY	2161	TTTCACAGGACATGGACAGAGATTTTAAAGAAATATTTTCACTTACCCCTACTTTC	2220
DB	2161	TTTCACAGGACATGGACAGAGATTTTAAAGAAATATTTTCACTTACCCCTACTTTC	2220
QY	2221	CAGGCATTCAAATTCAGATGCTGGTTCAGAGAGTTCAGTTACACCCGCCCTTG	2280
DB	2221	CAGGCATTCAAATTCAGATGCTGGTTCAGAGAGTTCAGTTACACCCGCCCTTG	2280
QY	2281	GATCAGCTCTCAGGGTTCAATGGGAATTCATTTCTGAAGATGAGATTTGAATATGAA	2340

DB	2281	GATCAGCTCTCAGGGTTCAATGGGAATTCATTTCTGAAGATGAGATTTGAATGAA	2340
QY	2341	RACCAGAGAGCTGGCAGAAAGAGAGGAGATTTCGAACGTGCTGACGTTTGAAGAC	2400
DB	2341	RACCAGAGAGCTGGCAGAAAGAGAGGAGATTTCGAACGTGCTGACGTTTGAAGAC	2400
QY	2401	CTCCTTTTGGCTTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAGTCTGT	2460
DB	2401	CTCCTTTTGGCTTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAGTCTGT	2460
QY	2461	GTCCACAGAGACTGGCAGGAGGATGTGTTGGTCAACCGGAAAGTGTGTAAGATC	2520
DB	2461	GTCCACAGAGACTGGCAGGAGGATGTGTTGGTCAACCGGAAAGTGTGTAAGATC	2520
QY	2521	TGTGACTTTGGACTGGCCGAGACATCTCAGCGACTCCAGCTACCTGTCAGGGGCAAC	2580
DB	2521	TGTGACTTTGGACTGGCCGAGACATCTCAGCGACTCCAGCTACCTGTCAGGGGCAAC	2580
QY	2581	GCACGGCTGCCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACAAATC	2640
DB	2581	GCACGGCTGCCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACAAATC	2640
QY	2641	AAGAGTACGCTGTGCTCTACGGCATCTCTCTCTGGGAGATATTTTCACTGGGTGAAAC	2700
DB	2641	AAGAGTACGCTGTGCTCTACGGCATCTCTCTCTGGGAGATATTTTCACTGGGTGAAAC	2700
QY	2701	CCTTACCTGGCATTCCTGTCAGCGTAACTTCTATAAATGATTCAGAGTGATTTAA	2760
DB	2701	CCTTACCTGGCATTCCTGTCAGCGTAACTTCTATAAATGATTCAGAGTGATTTAA	2760
QY	2761	ATGAGCAGCCATTCATGCCACAGAGGATATCTTTGTAATGCAATCTCTCTGGCT	2820
DB	2761	ATGAGCAGCCATTCATGCCACAGAGGATATCTTTGTAATGCAATCTCTCTGGCT	2820
QY	2821	TTTGTACTCAAGGAAGCGGCTATCTTCCCACACCTGACCTTCAATTTTAGATGTCAGCTG	2880
DB	2821	TTTGTACTCAAGGAAGCGGCTATCTTCCCACACCTGACCTTCAATTTTAGATGTCAGCTG	2880
QY	2881	GCAGGAGCAGAGAGCATGTATCAGACATCCATCCATCTACCAAAACAGCGGCCCT	2940
DB	2881	GCAGGAGCAGAGAGCATGTATCAGACATCCATCCATCTACCAAAACAGCGGCCCT	2940
QY	2941	CAGCAGAGAGCGGCTCAGAGCCAGTCGCCACAGCGCCAGTGAAGATTACACAGAA	3000
DB	2941	CAGCAGAGAGCGGCTCAGAGCCAGTCGCCACAGCGCCAGTGAAGATTACACAGAA	3000
QY	3001	AGAAATAGCAGAGAGGCTTGGACCCGCCACCTTAGCAGGCTGTAGACCGCAGAGCCA	3060
DB	3001	AGAAATAGCAGAGAGGCTTGGACCCGCCACCTTAGCAGGCTGTAGACCGCAGAGCCA	3060
QY	3061	AGATTAGCTCGCCTCTGAGGAGCGCCTACAGCGCTTCTTCCCTCGACTTTTCTCT	3120
DB	3061	AGATTAGCTCGCCTCTGAGGAGCGCCTACAGCGCTTCTTCCCTCGACTTTTCTCT	3120
QY	3121	AGATGCTGTGCCATTACTCCAAAGTGAATCTATAAATCAAACTCTCTCTCGCACAG	3180
DB	3121	AGATGCTGTGCCATTACTCCAAAGTGAATCTATAAATCAAACTCTCTCTCGCACAG	3180
QY	3181	CGCGGAGAGCCAAATTAATGAGACTTGTGTTGAGCGCGCTACCCCTGCGGCGCTTTCACG	3240
DB	3181	CGCGGAGAGCCAAATTAATGAGACTTGTGTTGAGCGCGCTACCCCTGCGGCGCTTTCACG	3240
QY	3241	AGCTTGGGGGAAACCATGTATCTGAATATAGTATATTTCTGTAATACGTGAACAA	3300
DB	3241	AGCTTGGGGGAAACCATGTATCTGAATATAGTATATTTCTGTAATACGTGAACAA	3300
QY	3301	ACCAACCCGTTTGTCTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTTAA	3360
DB	3301	ACCAACCCGTTTGTCTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTTAA	3360
QY	3361	AATACTGTAACCTTTTTCATCTATTATGTTGATATATTTTATGATGGAATAAATCTTC	3420
DB	3361	AATACTGTAACCTTTTTCATCTATTATGTTGATATATTTTATGATGGAATAAATCTTC	3420

Db	3361	AACTACTATGTAAGTATTTTCATCTATTATTAGTGATATATTTTATGGATGGAATAAACTTTC	3420
Qy	3421	TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA	3453
Db	3421	TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA	3453

## RESULT 6

AAQ35249  
ID AAQ35249 standard; cDNA; 3453 BP.

25-JUN-1993

Murine flk-2 coding sequence

Murine; receptor; protein; t

liver; spleen; thymus; adult

1. *Chlorophyll a* (Chl *a*)

—

CDS 31..3009

misc\_RNA

WO9300349-A.

07-JAN-1993.

26 - JUN - 1992; 92WO-US05401.

28-JUN-1991; 91US-0728913.

24-DEC-1991; 91US-0813593.

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818

Lemischka IR;  
WPI; 1993-036323/04.  
P-ESDB; AAR31375.

Nucleic acid encoding receptor protein tyrosine kinase - allows development of ligands to stimulate proliferation and/or

**Abstract**

СТАЛІ 3; РІЗ 1а; /opp; ENGLISH.

This sequence encodes a murine receptor protein tyrosine belongs to a new functional class of protein tyrosine

(CH<sub>3</sub>)<sub>2</sub>C=CH<sub>2</sub>

encoded by this sequence is an example of a receptor protein and is called fetal liver kinase 2 (flk-2). flk-2 is expressed in fetal liver

spleen and thymus, and adult brain and marrow occurs in the most primitive tunicate

to be uncommitted. Therefore, thymocytes expressing

be expres

sequence 3453 BP; 947 A; 821 C; 876 G; 809 T; 0 other;

Very Match 100.0%; score 3431.4

atches 3452; Conservative 0; Mismatches 1; Index

1081 ATAAACGCTACACGCTCGCAAGAGATGAAATGACCCGTACGAAAGTTCTGCTC 1140  
1141 TCAGTCAGGTTTAAAGCGTACCCAGATCCGATCGACGTGATCTTCTCTCAAGCCTCA 1200  
1141 TCAGTCAGGTTTAAAGCGTACCCAGATCCGATCCGATCGACGTGATCTTCTCTCAAGCCTCA 1200  
1201 TTTCTCTGCTGAACAGAGAGGCTCGAGGATGGGTACAGCATATCTAAATTTTTCGATCAT 1260  
1201 TTTCTCTGCTGAACAGAGAGGCTCGAGGATGGGTACAGCATATCTAAATTTTTCGATCAT 1260  
1261 AAGAACAGCCAGGAGATACATATCTATGAGAAATGATGACGCCAGTTTCAACCAA 1320  
1321 ATGTTACGCTGAATATAGAAGAAACCTCAAGTGTAGCAAAATGCCCTACGCCAGCCAG 1380  
1321 ATGTTACGCTGAATATAGAAGAAACCTCAAGTGTAGCAAAATGCCCTACGCCAGCCAG 1380  
1381 CGGTCCTGCTCTGTATGGCTACCGCTACCCCTTGGACCTGGAAGAGTGTTCGGAC 1440  
1381 CGGTCCTGCTCTGTATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTGTTCGGAC 1440  
1441 AAATCTCCAAATTCACGAGGAAATCCAGAGGAGTTTGAATATAAAAGGCTAACACA 1500  
1441 AAATCTGCCAATTCACGAGGAAATCCAGAGGAGTTTGAATATAAAAGGCTAACACA 1500  
1501 AAAGTGTGGCCAGTGGGTGTCAGCAGTACTCTAAATATAGTGAAGCCGGGAAGGG 1560  
1501 AAAGTGTGGCCAGTGGGTGTCAGCAGTACTCTAAATATAGTGAAGCCGGGAAGGG 1560  
1561 CTCTGGTCAATGCTGTGCTGACAAATCTATGGCAGCTCTGGAAACCATCTTTTA 1620  
1561 CTCTGGTCAATGCTGTGCTGACAAATCTATGGCAGCTCTGGAAACCATCTTTTA 1620  
1621 AACTCACCAGGCCCTTCCCTTTCATCCAAGACACATCTCTTATGCGACCATTTGG 1680  
1621 AACTCACCAGGCCCTTCCCTTTCATCCAAGACACATCTCTTATGCGACCATTTGG 1680  
1681 CTCTGCTCCGCTTCATTTGTTGTTCTCATTTGTTGATGTCGCCAATAACAAAGCAA 1740  
1681 CTCTGCTCCGCTTCATTTGTTGTTCTCATTTGTTGATGTCGCCAATAACAAAGCAA 1740  
1741 TTTAGGTACGAGTACGCTGATGATCCAGGTGACTGGCCCTGTGATACAGGATAC 1800  
1741 TTTAGGTACGAGTACGCTGATGATCCAGGTGACTGGCCCTGTGATACAGGATAC 1800  
1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAAC 1860  
1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAAC 1860  
1861 TTAGAGTTGGAGAGGCTCGGGGCTGCGGCTTTCCGGAGGGTGATGAACGCCAGGCC 1920  
1861 TTAGAGTTGGAGAGGCTCGGGGCTGCGGCTTTCCGGAGGGTGATGAACGCCAGGCC 1920  
1921 TATGCAATAGTAAACGGAGCTCAATTCAGGTGGCGTGAAGTGTCTAAAGAGAA 1980  
1921 TATGCAATAGTAAACGGAGCTCAATTCAGGTGGCGTGAAGTGTCTAAAGAGAA 1980  
1981 GCTGACAGCTGTGAAAAGAAAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040  
1981 GCTGACAGCTGTGAAAAGAAAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040  
2041 CACCATGAACATCGTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTGTACTTG 2100  
2041 CACCATGAACATCGTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTGTACTTG 2100  
2101 ATTTTGAATATGTTGCTATGGTGACCTTCACTACCTTGAAGTAAAGAGAGAAG 2160  
2101 ATTTTGAATATGTTGCTATGGTGACCTTCACTACCTTGAAGTAAAGAGAGAAG 2160  
2161 TTTACAGGACATGACAGAGATTTTGAAGAACATATTTTCTTCTTACCTACTTTC 2220

2161 TTTACAGGACATGGACAGAGATTTTAAAGAACATAATTTTACGTTCTTACCCTACTTTC 2220  
2221 CAGGACATTTCAATTTCCAGCATGCTGTTTCCAGGAAAGTTTACAGTACACCCGCTTG 2280  
2221 CAGGACATTTCAATTTCCAGCATGCTGTTTCCAGGAAAGTTTACAGTACACCCGCTTG 2280  
2281 GATCAGCTCTCAGGGTTCAATGGGAATCAATTCATTTCTCAAGATGAGATTGATGAA 2340  
2281 GATCAGCTCTCAGGGTTCAATGGGAATCAATTCATTTCTCAAGATGAGATTGATGAA 2340  
2341 AACCAAGAGGCTGCGAGAAAGAGAGAGATTTTGAACGCTGCTGACGTTTGAAGAC 2400  
2341 AACCAAGAGGCTGCGAGAAAGAGAGAGATTTTGAACGCTGCTGACGTTTGAAGAC 2400  
2401 CTCCCTTTGCTTTTCCGTACCAAGTGGCCAAAGGATGGAATCTCTGGAGTTCAAGTCTGT 2460  
2401 CTCCCTTTGCTTTTCCGTACCAAGTGGCCAAAGGATGGAATCTCTGGAGTTCAAGTCTGT 2460  
2461 GTCCACAGAGACTGCGCAGCCAGGAATGTTTGGTCACCCACCGGAAGTGGTCAAGATC 2520  
2461 GTCCACAGAGACTGCGCAGCCAGGAATGTTTGGTCACCCACCGGAAGTGGTCAAGATC 2520  
2521 TGTGACTTTGGACTGGCCCGAGACATCTCTGAGCGACTCCAGTACGTCGTCAGGGGCAAC 2580  
2521 TGTGACTTTGGACTGGCCCGAGACATCTCTGAGCGACTCCAGTACGTCGTCAGGGGCAAC 2580  
2581 GCACGGCTCCGGTGAAGTGGATGGCACCGGAGAGCTTATTTGAAGGGATCTACACAATC 2640  
2581 GCACGGCTCCGGTGAAGTGGATGGCACCGGAGAGCTTATTTGAAGGGATCTACACAATC 2640  
2641 AAGAGTACGCTCTGCTGCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGCAAC 2700  
2641 AAGAGTACGCTCTGCTGCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGCAAC 2700  
2701 CCTTACCTCGCATTCCTGTCGACGCTAATCTTCTATAAAGTATTCAGAGTGGATTTAA 2760  
2701 CCTTACCTCGCATTCCTGTCGACGCTAATCTTCTATAAAGTATTCAGAGTGGATTTAA 2760  
2761 ATGGAGCAGCATTCATGCCACAGAGGATATCTTGTATGCAATCTCTCTGCTGGCT 2820  
2761 ATGGAGCAGCATTCATGCCACAGAGGATATCTTGTATGCAATCTCTCTGCTGGCT 2820  
2821 TTTGACTCAAGGAAGCGGCATCTTCTCCCACTGACTTCTTATAGGATGTCAGCTG 2880  
2821 TTTGACTCAAGGAAGCGGCATCTTCTCCCACTGACTTCTTATAGGATGTCAGCTG 2880  
2881 GCAGAGGCAAGAAAGCATGTATCAGAACATCCATCCATCTACCAAAACAGGCGGCCCT 2940  
2881 GCAGAGGCAAGAAAGCATGTATCAGAACATCCATCCATCTACCAAAACAGGCGGCCCT 2940  
2941 CAGCAGAGGCGGGCTCAGAGCCAGTCCGACAGCGCCAGCGGAGTGAAGATTACAGAGAA 3000  
2941 CAGCAGAGGCGGGCTCAGAGCCAGTCCGACAGCGCCAGCGGAGTGAAGATTACAGAGAA 3000  
3001 AGAAGTTAGCGAGAGGCTTTGGACCCCGCCAGCTAGCAGGCTGTAGCCGAGAGCA 3060  
3001 AGAAGTTAGCGAGAGGCTTTGGACCCCGCCAGCTAGCAGGCTGTAGCCGAGAGCA 3060  
3061 AGATTAGCTCGCTCTGAGGAAGCGCCCTACAGCGCTTGTCTGCTGAGCATTTTCTCT 3120  
3061 AGATTAGCTCGCTCTGAGGAAGCGCCCTACAGCGCTTGTCTGCTGAGCATTTTCTCT 3120  
3121 AGATGCTGTGCTTACTTCCAAAGTGACTTCTATAAAATCAAACTCTCTCCGACAG 3180  
3121 AGATGCTGTGCTTACTTCCAAAGTGACTTCTATAAAATCAAACTCTCTCCGACAG 3180  
3181 GCGGGAGAGCCATATGAGACTTGTGGTGACCCGCCCTACCTGCGGGGCTTTCCACG 3240  
3181 GCGGGAGAGCCATATGAGACTTGTGGTGACCCGCCCTACCTGCGGGGCTTTCCACG 3240  
3241 AGCTTCAGGGGAAAGCCATGTATCTGAATATAGTATATTTCTTGTAAATACGTGAACAA 3300  
3241 AGCTTCAGGGGAAAGCCATGTATCTGAATATAGTATATTTCTTGTAAATACGTGAACAA 3300

QY 3301 ACCAACCCGTTTTTGTCTAAGCGAAAGCTAAATATGATTTTAAAAATCTATGTTTTAA 3360  
|||||  
Db 3301 ACCAACCCGTTTTTGTCTAAGCGAAAGCTAAATATGATTTTAAAAATCTATGTTTTAA 3360  
QY 3361 AATACTATGTAACCTTTTTCATCTATTTAGTGATATATTTATGGATGGAATAAACTTTC 3420  
|||||  
Db 3361 AATACTATGTAACCTTTTTCATCTATTTAGTGATATATTTATGGATGGAATAAACTTTC 3420  
QY 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453  
|||||  
Db 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453

RESULT 7  
AAQ79068  
ID AAQ79068 standard; cDNA: 3453 BP.  
XX  
AC AAQ79068;  
XX 04-JUL-1995 (first entry)  
XX Mouse flk-2 cDNA.  
XX Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;  
KW hematopoiesis; stem cell; ds.  
XX  
OS Mus sp.

XX Key Location/Qualifiers  
FH 58..3039  
CDS /\*tag= a  
FT sig\_peptide 58..138  
FT /\*tag= b  
FT mat\_peptide 139..3036  
FT /\*tag= c  
XX US5367057-A.

XX 22-NOV-1994.  
XX  
XX 02-APR-1991; 91US-0679666.  
XX  
XX 02-APR-1991; 91US-0679666.  
PR 28-JUN-1991; 91US-0728913.  
PR 15-NOV-1991; 91US-0793065.  
PR 24-DEC-1991; 91US-0813593.  
PR 26-JUN-1992; 92US-0906397.  
PR 12-NOV-1992; 92US-0975049.  
PR 19-NOV-1992; 92US-0977451.  
PR 30-APR-1993; 93US-0055269.  
XX  
XX (UYPR-) UNIV PRINCETON.  
XX  
XX Lemischka IR;  
XX  
XX WPI: 1995-005894/01.  
DR P-PSDB; AAR67535.  
XX  
XX Murine flk-2 receptor protein tyrosine kinase - used to stimulate  
PT proliferation and/or stimulation of primitive mammalian  
PT haematopoietic stem cells in vitro or in vivo.  
XX  
XX Disclosure; Fig. 1A-1F; 69pp; English.  
XX  
XX cDNAs encoding receptor protein tyrosine-kinases; mouse fetal liver  
CC kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAQ79068-70,  
CC respectively, and the deduced amino acid sequences in AAR67535-37,  
CC respectively.  
XX  
XX Sequence 3453 BP; 946 A; 822 C; 876 G; 809 T; 0 other;  
SQ  
Query Match 100.0%; Score 3451.4; DB 16; Length 3453;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3452; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GCGGCTGCTACCGCGCTCCGGAGGCCATCGGGCGTTGGCGAGCGCAGCGACCGG 60  
|||||  
Db 1 GCGGCTGCTACCGCGCTCCGGAGGCCATCGGGCGTTGGCGAGCGCAGCGACCGG 60  
QY 61 CGGCTGCTCTGCTGCTTTGTTGTCAGTAATGATTTCTTGAGACCGTTACAAACCAAGAC 120  
|||||  
Db 61 CGGCTGCTCTGCTGCTTTGTTGTCAGTAATGATTTCTTGAGACCGTTACAAACCAAGAC 120  
QY 121 CTGCTGTGATCAAGTGTCTTTTAAATAGTATGAGAACAAATGGCTCATCAGCGGGAAG 180  
|||||  
Db 121 CTGCTGTGATCAAGTGTCTTTTAAATAGTATGAGAACAAATGGCTCATCAGCGGGAAG 180  
QY 181 CCATCATCTAGCGAATGTCGAGGATCCCCAGAGACCTCCAGTGTACCCCGAGGCG 240  
|||||  
Db 181 CCATCATCTAGCGAATGTCGAGGATCCCCAGAGACCTCCAGTGTACCCCGAGGCG 240  
QY 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGAGGTGGCCGAGTCTGGTCCATC 300  
|||||  
Db 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGAGGTGGCCGAGTCTGGTCCATC 300  
QY 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCTGCTCTGGTCTTTAAGCAC 360  
|||||  
Db 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGACCTTTCTGCTCTGGTCTTTAAGCAC 360  
QY 361 AGCTCCCTGGCTGCCAGCGGCACCTTTGATTTACAAACAGAGAGAAATCTTCATGSCC 420  
|||||  
Db 361 AGCTCCCTGGCTGCCAGCGGCACCTTTGATTTACAAACAGAGAGAAATCTTCATGSCC 420  
QY 421 ATCTTGAAGTGCAGAGACCCAGGAGAGATACTACTCCATATTCAGAGCGAAGCG 480  
|||||  
Db 421 ATCTTGAAGTGCAGAGACCCAGGAGAGATACTACTCCATATTCAGAGCGAAGCG 480  
QY 481 GCCAATACACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGCTAAGG 540  
|||||  
Db 481 GCCAATACACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGCTAAGG 540  
QY 541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCACCTCTCCATCTCCGAGGGTGT 600  
|||||  
Db 541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCACCTCTCCATCTCCGAGGGTGT 600  
QY 601 CCGGAGCCCACTGTGGAGTGGGTGCTCTGCAGCTCCACAGGAGAAAGCTGTAAAGAGAA 660  
|||||  
Db 601 CCGGAGCCCACTGTGGAGTGGGTGCTCTGCAGCTCCACAGGAGAAAGCTGTAAAGAGAA 660  
QY 661 GGCCCTGCTGTTGTCAGAAAGAGGAAAGGTACTTCAATGAGTTGTTCCGAAACAGACATC 720  
|||||  
Db 661 GGCCCTGCTGTTGTCAGAAAGAGGAAAGGTACTTCAATGAGTTGTTCCGAAACAGACATC 720  
QY 721 AGATGCTGTCTAGAAATGCACTGGCCGCGAATGCACCAAGCTGTTTCCACATAGATCTA 780  
|||||  
Db 721 AGATGCTGTCTAGAAATGCACTGGCCGCGAATGCACCAAGCTGTTTCCACATAGATCTA 780  
QY 781 AACAGGCTCTCTCAGAGCACACTGCCCACTTTTCTGAAAGTGGGGGAAACCTTGTGG 840  
|||||  
Db 781 AACAGGCTCTCTCAGAGCACACTGCCCACTTTTCTGAAAGTGGGGGAAACCTTGTGG 840  
QY 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGATTCGGGCTCACTCCAGAGCTGGAGAC 900  
|||||  
Db 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGATTCGGGCTCACTCCAGAGCTGGAGAC 900  
QY 901 AAGCCCTGGAGAGGCGAGCTACTTTGAGATGAGTACCTACTCTCCACAAACAGACCATG 960  
|||||  
Db 901 AAGCCCTGGAGAGGCGAGCTACTTTGAGATGAGTACCTACTCTCCACAAACAGACCATG 960  
QY 961 ATTCGATTCCTCTGCGCTTTTGTCTTCGTTGGGAAGAACACACCGGATATTACACC 1020  
|||||  
Db 961 ATTCGATTCCTCTGCGCTTTTGTCTTCGTTGGGAAGAACACACCGGATATTACACC 1020  
QY 1021 TGCTCTTCTCTCAAGCACCCAGCCAGCTAGCGTTGCTGACCATCTCTAGAAAAAGGTTT 1080  
|||||

Db 1021 TGCTCTTCTCAAAAGCACCCAGCAGCTCAGCGTTGGTGACCATCTAGAAAAAGGGTTT 1080  
Qy 1081 ATAAACGCTACCACTCGCAGAGAGATATGAATTTGACCCGTACGAAAGATTCTGCTTC 1140  
Db 1081 ATAAACGCTACCACTCGCAGAGAGATATGAATTTGACCCGTACGAAAGATTCTGCTTC 1140  
Qy 1141 TCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGCGATGCTCTCTCAAGCCCTCA 1200  
Db 1141 TCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGCGATGCTCTCTCAAGCCCTCA 1200  
Qy 1201 TTTCCTTGTAACAGAGAGCGCTGGAGGTGAGGATACAGCATATCTAAATTTTGGATCAT 1260  
Db 1201 TTTCCTTGTAACAGAGAGCGCTGGAGGTGAGGATACAGCATATCTAAATTTTGGATCAT 1260  
Qy 1261 AAGAACAAGCAGGAGAGTACATATCTATCGACAAATATGACGCCAGCTGTCACCAAA 1320  
Db 1261 AAGNACAAGCAGGAGAGTACATATCTATCGACAAATATGACGCCAGCTGTCACCAAA 1320  
Qy 1321 ATGTTCAAGCTGTAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380  
Db 1321 ATGTTCAAGCTGTAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380  
Qy 1381 GCGTCTGTTCTCTGATGGCTACCCGTACCCGTACCCGTGAGGACTGGAAGAGTGTTCGGAC 1440  
Db 1381 GCGTCTGTTCTCTGATGGCTACCCGTACCCGTACCCGTGAGGACTGGAAGAGTGTTCGGAC 1440  
Qy 1441 AAATCTCCCAATTCACGAGGAATCCAGAGAGTGTGGATATAAAGGCTTAACAGA 1500  
Db 1441 AAATCTCCCAATTCACGAGGAATCCAGAGAGTGTGGATATAAAGGCTTAACAGA 1500  
Qy 1501 AAAGTGTGTTGCCAGTGGGTGTCAGCAGTACTCTAAATATGAGTGAGGCGCGGAAAGGG 1560  
Db 1501 AAAGTGTGTTGCCAGTGGGTGTCAGCAGTACTCTAAATATGAGTGAGGCGCGGAAAGGG 1560  
Qy 1561 CTCTGCTGTAATCTGTGCGTACAAATCTATGAGGACGCTTTCGCAAAACATCTTTTA 1620  
Db 1561 CTCTGCTGTAATCTGTGCGTACAAATCTATGAGGACGCTTTCGCAAAACATCTTTTA 1620  
Qy 1621 AACTCAGCAGCCCTTCCCTTTCATCCAGACAACTCTCTTCTATGCGACCATGGG 1680  
Db 1621 AACTCAGCAGCCCTTCCCTTTCATCCAGACAACTCTCTTCTATGCGACCATGGG 1680  
Qy 1681 CTCTGCTCCTCTCATTTGTTCTCATTTGTTGATCTGCCCAATAACAAAAGCAA 1740  
Db 1681 CTCTGCTCCTCTCATTTGTTCTCATTTGTTGATCTGCCCAATAACAAAAGCAA 1740  
Qy 1741 TTTAGTACGAGTACGTCAGATGATCAGGTGACTGCGCCCTCGGATACAGAGTAC 1800  
Db 1741 TTTAGTACGAGTACGTCAGATGATCAGGTGACTGCGCCCTCGGATACAGAGTAC 1800  
Qy 1801 TTCTACGTTCACTCAGGACTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAAC 1860  
Db 1801 TTCTACGTTCACTCAGGACTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAAC 1860  
Qy 1861 TTAGAGTTTGGAGGTCCTGGGGTCTGGCGCTTTCGGAGGTTGATGAACCCACGGCC 1920  
Db 1861 TTAGAGTTTGGAGGTCCTGGGGTCTGGCGCTTTCGGAGGTTGATGAACCCACGGCC 1920  
Qy 1921 TATGGCATTAAGTAAACGGGAGTCTCAATTCAGTGGCGGTGAGATGCTTAAAGAGAAA 1980  
Db 1921 TATGGCATTAAGTAAACGGGAGTCTCAATTCAGTGGCGGTGAGATGCTTAAAGAGAAA 1980  
Qy 1981 GCTGACAGCTGTGAAAAGAGTCTCATGTCGAGGCTCAAAATGATGACCCCTGGGA 2040  
Db 1981 GCTGACAGCTGTGAAAAGAGTCTCATGTCGAGGCTCAAAATGATGACCCCTGGGA 2040  
Qy 2041 CACCATCACAACTGCGTAATCTGCTGGGSCATGCACACTGTCAGGSCCAGTGTACTTG 2100  
Db 2041 CACCATCACAACTGCGTAATCTGCTGGGSCATGCACACTGTCAGGSCCAGTGTACTTG 2100  
Qy 2101 ATTTTGAATATGTTGCTATGTTGAGCTCTCCTCACTACCTAAGAAGTAAAGAGAGAG 2160  
Db 2101 ATTTTGAATATGTTGCTATGTTGAGCTCTCCTCACTACCTAAGAAGTAAAGAGAGAG 2160

Qy 2161 TTTACAGAGATGACAGAGATTTTAAAGAACATAATTTTCAGTTCCTTACCTACTTTC 2220  
Db 2161 TTTACAGAGATGACAGAGATTTTAAAGAACATAATTTTCAGTTCCTTACCTACTTTC 2220  
Qy 2221 CAGGACATTTCAAATTTCCAGCATGCTGTTTACAGAGAGTTCAGTTACACCCGCTTTC 2280  
Db 2221 CAGGACATTTCAAATTTCCAGCATGCTGTTTACAGAGAGTTCAGTTACACCCGCTTTC 2280  
Qy 2281 GATCAGCTCTCAGGTTTCAATGGGAATTCATTTCTGAAGATGAGATTTGAATATGAA 2340  
Db 2281 GATCAGCTCTCAGGTTTCAATGGGAATTCATTTCTGAAGATGAGATTTGAATATGAA 2340  
Qy 2341 AACCAAGAGAGCTGGCAGAGAGAGGAGGAGATTTGAACGCTGCTGAGCTTGAAGAC 2400  
Db 2341 AACCAAGAGAGCTGGCAGAGAGAGGAGGAGATTTGAACGCTGCTGAGCTTGAAGAC 2400  
Qy 2401 CTCTTTTCTTTCCTTACCAAGTGGCCAAAGCATGGAAATTCCTGGAGTTCAGTCTGT 2460  
Db 2401 CTCTTTTCTTTCCTTACCAAGTGGCCAAAGCATGGAAATTCCTGGAGTTCAGTCTGT 2460  
Qy 2461 GTCCACAGAGCTGGCAGCAGAGGATGTTGGTACCCACCGGAGAGTGTGAAGATC 2520  
Db 2461 GTCCACAGAGCTGGCAGCAGGAGATGTTGGTACCCACCGGAGAGTGTGAAGATC 2520  
Qy 2521 TGTGACTTTGGACTGGCCGAGACATCCTGAGGACTCCAGCTACGTCGTCAGGGCAAC 2580  
Db 2521 TGTGACTTTGGACTGGCCGAGACATCCTGAGGACTCCAGCTACGTCGTCAGGGCAAC 2580  
Qy 2581 GCACGGCTGCGGTTGAAGTGGATGGCACCAGAGAGCTTATTTGAAGGATCTACACAATC 2640  
Db 2581 GCACGGCTGCGGTTGAAGTGGATGGCACCAGAGAGCTTATTTGAAGGATCTACACAATC 2640  
Qy 2641 AAGAGTACGCTGCTCTACGGCATCTCTCTCTGGAGATATTTTCACTGGGTGGAAC 2700  
Db 2641 AAGAGTACGCTGCTCTACGGCATCTCTCTCTGGAGATATTTTCACTGGGTGGAAC 2700  
Qy 2701 CCTTACCTGGCATCTCTGTCAGCCTTCTATTAATGATTCAGAGTGGATTTAA 2760  
Db 2701 CCTTACCTGGCATCTCTGTCAGCCTTCTATTAATGATTCAGAGTGGATTTAA 2760  
Qy 2761 ATGAGCAGCCTTCTATGCCACAGAGGATATCTTTGTAATGCAATCTCTCTGGCT 2820  
Db 2761 ATGAGCAGCCTTCTATGCCACAGAGGATATCTTTGTAATGCAATCTCTCTGGCT 2820  
Qy 2821 TTTGACTCAAGAGAGCGGCTTCCCTTCCCAACCTGACTTCATTTTAGGATGTCAGCTG 2880  
Db 2821 TTTGACTCAAGAGAGCGGCTTCCCTTCCCAACCTGACTTCATTTTAGGATGTCAGCTG 2880  
Qy 2881 GCAGAGGAGAGAGATGATCAGACATCCATCCATCTACCAAAACAGCGGCCCCCT 2940  
Db 2881 GCAGAGGAGAGAGATGATCAGACATCCATCCATCTACCAAAACAGCGGCCCCCT 2940  
Qy 2941 CAGCAGAGAGCGGCTCAGAGCCAGTCGCCACAGCGCAGGTGAAGATTCACAGAGAA 3000  
Db 2941 CAGCAGAGAGCGGCTCAGAGCCAGTCGCCACAGCGCAGGTGAAGATTCACAGAGAA 3000  
Qy 3001 AGAAGTTAGCAGAGGAGGCTTTGGACCCCGCCACCTAGCAGCTGTAGACCCGAGAGCCA 3060  
Db 3001 AGAAGTTAGCAGAGAGGAGGCTTTGGACCCCGCCACCTAGCAGCTGTAGACCCGAGAGCCA 3060  
Qy 3061 AGATTAGCTCTGCTCTGAGGAGGCGCTTACAGCGCTTGTCTGCTGGAGCTTTCTCT 3120  
Db 3061 AGATTAGCTCTGCTCTGAGGAGGCGCTTACAGCGCTTGTCTGCTGGAGCTTTCTCT 3120  
Qy 3121 AGATGCTCTGCTCTTACTCCAAAGTGAATTTCTATAAATCAAACTCTCTCGCACAG 3180  
Db 3121 AGATGCTCTGCTCTTACTCCAAAGTGAATTTCTATAAATCAAACTCTCTCGCACAG 3180  
Qy 3181 GCGGAGAGGCAATTAATGAGCTTTGTTGAGCGCGCTACCTCGGGGCTTTTCCAG 3240  
Db 3181 GCGGAGAGGCAATTAATGAGCTTTGTTGAGCGCGCTACCTCGGGGCTTTTCCAG 3240

QY 3241 AGCTTGAGGGGAAAGCCATGATCTGAAATATAGTATATATCTTGTAATAGCTGAACAA 3300  
 Db |||||||  
 3241 AGCTTGAGGGGAAAGCCATGATCTGAAATATAGTATATATCTTGTAATAGCTGAACAA 3300  
 QY 3301 ACCAAACCCGTTTTTGTGTAAGGGAAGCTAAATATGATTTTAAATATCTATGTTTTAA 3360  
 Db |||||||  
 3301 ACCAAACCCGTTTTTGTGTAAGGGAAGCTAAATATGATTTTAAATATCTATGTTTTAA 3360  
 QY 3361 AATACTATGTAACCTTTTTCATCTATTTAGTGATATATTTTGGATGGAATAAACTTTC 3420  
 Db |||||||  
 3361 AATACTATGTAACCTTTTTCATCTATTTAGTGATATATTTTGGATGGAATAAACTTTC 3420  
 QY 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAA 3453  
 Db |||||||  
 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAA 3453

## RESULT 8

AAQ29954  
 ID AAQ29954 standard; cDNA; 3453 BP.  
 AC AAQ29954;  
 XX  
 DT 15-MAR-1993 (first entry)  
 XX  
 DE Murine flk-2 cDNA sequence.  
 XX  
 KW Thymidine kinase; TK; haematopoietic; stem cells; proliferation;  
 KW differentiation; progenitor cells; foetal liver kinase; ss.  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 31..3009  
 FT /\*tag= a  
 XX  
 PN W09217486-A.  
 XX  
 PD 15-OCT-1992.  
 XX  
 PF 02-APR-1992; 92WO-US02750.  
 XX  
 PR 02-APR-1991; 91US-0679666.  
 PR 28-JUN-1991; 91US-0728913.  
 PR 15-NOV-1991; 91US-0793065.  
 PR 24-DEC-1991; 91US-0813593.  
 XX  
 PA (UYPR-) UNIV PRINCETON.  
 XX  
 PI Lemischka IR;  
 XX  
 DR WPI; 1992-366185/44.  
 DR P-PSDB; AAK28038.  
 XX  
 PT Stimulating proliferation and/or differentiation of primitive  
 PT mammalian haematopoietic stem cells - using ligand that binds  
 PT thymidine kinase and flk-1 and flk-2  
 XX  
 PS Claim 6; Fig 1a; 94pp; English.  
 XX  
 CC The murine foetal liver kinase (flk)-2 clone was isolated by  
 CC standard PCR techniques from stem-cell receptor-contg. tissue cDNA  
 CC libraries. Suitable tissues include foetal liver, spleen or thymus  
 CC cells or adult marrow or brain cells. The PCR primers used are based  
 CC on known sections of the flk-2 gene. The murine flk-2 clone may be  
 CC used in a vector to transform haematopoietic cells. The thymidine  
 CC kinase encoded by flk-2 is expressed in primitive but not mature  
 CC haematopoietic cells. Ligand binding to the TK may be prep'd. which  
 CC can stimulate proliferation and/or differentiation of primitive  
 CC haematopoietic cells in vivo. The ligands can stimulate the  
 CC proliferation of additional primitive stem cells, differentiation into  
 CC more mature progenitor cells, or both.  
 CC See also AAQ29955-7.

XX Sequence 3453 BP; 946 A; 822 C; 876 G; 809 T; 0 other;  
 SQ  
 Query Match 99.9%; Score 3448.2; DB 13; Length 3453;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 3450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 GCGGCTGGCTACCGCGGCTCCGGAGGCCATCGCGGCGTTGGCGAGCGACGACCGG 60  
 Db |||||||  
 1 GCGGCTGGCTACCGCGGCTCCGGAGGCCATCGCGGCGTTGGCGAGCGACGACCGG 60  
 QY 61 GCGCTGCTGCTGCTGTTTGTGTCAGTAATGATTTTGAGACCGTTTACAACCAAGAC 120  
 Db |||||||  
 61 GCGCTGCTGCTGCTGTTTGTGTCAGTAATGATTTTGAGACCGTTTACAACCAAGAC 120  
 QY 121 CTGCTGCTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGAAAG 180  
 Db |||||||  
 121 CTGCTGCTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGAAAG 180  
 QY 181 CCATCATCGTACCGAATGTGCGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGCGC 240  
 Db |||||||  
 181 CCATCATCGTACCGAATGTGCGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGCGC 240  
 QY 241 CAGAGTGAAGGACGCGTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCTGGTCCATC 300  
 Db |||||||  
 241 CAGAGTGAAGGACGCGTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCTGGTCCATC 300  
 QY 301 ACCCTGCAAGTGCAGTCCGCCACCCAGGGGACCTTTCTGCTCTGGTCTTTAAGCAC 360  
 Db |||||||  
 301 ACCCTGCAAGTGCAGTCCGCCACCCAGGGGACCTTTCTGCTCTGGTCTTTAAGCAC 360  
 QY 361 AGCTCCCTGGCTGCCAGCGCACCTTTGATTTACAAAACAGAGAGATCGTTCCATGGCC 420  
 Db |||||||  
 361 AGCTCCCTGGCTGCCAGCGCACCTTTGATTTACAAAACAGAGAGATCGTTCCATGGCC 420  
 QY 421 ATCTTGAACGTGACAGAGACCCAGGAGAGATACCTACTCTATATTCAGAGCGAACGC 480  
 Db |||||||  
 421 ATCTTGAACGTGACAGAGACCCAGGAGAGATACCTACTCTATATTCAGAGCGAACGC 480  
 QY 481 GCCAACTACACAGTACTGTTTCACAGTGAATGTAAAGATACACAGCTGTATGCTAAGG 540  
 Db |||||||  
 481 GCCAACTACACAGTACTGTTTCACAGTGAATGTAAAGATACACAGCTGTATGCTAAGG 540  
 QY 541 AGACCTTACTTTAGGAAGATGGAACACAGAGTGCACCTGCTGCACTCTCCGAGGCTTT 600  
 Db |||||||  
 541 AGACCTTACTTTAGGAAGATGGAACACAGAGTGCACCTGCTGCACTCTCCGAGGCTTT 600  
 QY 601 CCGGAGCCCACTGTGGAGTGGGTGCTCTGAGCTCCACAGGAAAGCTGTAAGAAGAA 660  
 Db |||||||  
 601 CCGGAGCCCACTGTGGAGTGGGTGCTCTGAGCTCCACAGGAAAGCTGTAAGAAGAA 660  
 QY 661 GGCCCTGCTGTTGTCAGAAAGGAGGAAAGGTACTTTCATGAGTGTTCGGAACAGACATC 720  
 Db |||||||  
 661 GGCCCTGCTGTTGTCAGAAAGGAGGAAAGGTACTTTCATGAGTGTTCGGAACAGACATC 720  
 QY 721 AGATGCTGTGCTAGAAATGCACCTGGGCGCGGAATGCACCAAGCTGTTTCAACATAGATCTA 780  
 Db |||||||  
 721 AGATGCTGTGCTAGAAATGCACCTGGGCGCGGAATGCACCAAGCTGTTTCAACATAGATCTA 780  
 QY 781 AACCAGGCTCCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGGAACCCCTTGTGG 840  
 Db |||||||  
 781 AACCAGGCTCCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGGAACCCCTTGTGG 840  
 QY 841 ATCAGTGTAAAGCCCATCATGTGAACCATGGATTTCGGGCTCACCCTGGGAGCTGGAAGAC 900  
 Db |||||||  
 841 ATCAGTGTAAAGCCCATCATGTGAACCATGGATTTCGGGCTCACCCTGGGAGCTGGAAGAC 900  
 QY 901 AAGCCCTCGGAGGCGCAGCTACTTTGAGATGAGTACCTACTTCCACAAACAGCACCATG 960  
 Db |||||||  
 901 AAGCCCTCGGAGGCGCAGCTACTTTGAGATGAGTACCTACTTCCACAAACAGCACCATG 960  
 QY 961 ATTCCGATCTCTTTGGCCCTTTGTGCTCTTCGTTGGGAGGAACGACACCCGGATTATACCC 1020  
 Db |||||||

Db 961 AFTCGGATTCTCTTGGCCCTTTGTGCTCTCCGTGGGAAGAACGACACCGGATATTACACC 1020  
Qy 1021 TGCTCTTCTCTAAAGCACCCAGCCAGTCAGCGTTGGTGACCATCTCTAGAAAAGGGTTT 1080  
Db 1021 TGCTCTTCTCTAAAGCACCCAGCCAGTCAGCGTTGGTGACCATCTCTAGAAAAGGGTTT 1080  
Qy 1081 ATAAACGCTACCGCTCGCAAGAAGAGTATGAATTTGACCCGTAGCAAAAGTTCTGCTTC 1140  
Db 1081 ATAAACGCTACCGCTCGCAAGAAGATATGAATTTGACCCGTACGAAAAGTTCTGCTTC 1140  
Qy 1141 TCAGTCAGTTTAAACGCTACCCAGCAATCCGATCGACGTGATCTTCTCTCAAGCCTCA 1200  
Db 1141 TCAGTCAGTTTAAACGCTACCCAGCAATCCGATCGACGTGATCTTCTCTCAAGCCTCA 1200  
Qy 1201 TTTCTCTGTGAACAGAGGCTCGAGGATGGGTACAGATATCTAAATTTTGGCATCAT 1260  
Db 1201 TTTCTCTGTGAACAGAGGCTCGAGGATGGGTACAGATATCTAAATTTTGGCATCAT 1260  
Qy 1261 AAGAACAAAGCCAGGAGGTACATATTCTATGCAGAAAATGATGAGCCCGAGTTTCAACAAA 1320  
Db 1261 AAGAACAAAGCCAGGAGGTACATATTCTATGCAGAAAATGATGAGCCCGAGTTTCAACAAA 1320  
Qy 1321 ATGTTTCAGCTGAATATAGAAGAAACCTCAAGTGTAGCAAAATGCCCTAGCCAGCCAG 1380  
Db 1321 ATGTTTCAGCTGAATATAGAAGAAACCTCAAGTGTAGCAAAATGCCCTAGCCAGCCAG 1380  
Qy 1381 GCGTCCTGTCTCTGATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTTGTCGGAC 1440  
Db 1381 GCGTCCTGTCTCTGATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTTGTCGGAC 1440  
Qy 1441 AAATCTCCCAATTGCACGAGGAAATCCAGAAAGAGTTTGGAAATAAAAAGGCTTAACAGA 1500  
Db 1441 AAATCTCCCAATTGCACGAGGAAATCCAGAAAGAGTTTGGAAATAAAAAGGCTTAACAGA 1500  
Qy 1501 AAAGTGTGTTGGCCAGTGGGTGTCAGCAGTACTCTAAATATAGTGAAGCCGGGAAGGG 1560  
Db 1501 AAAGTGTGTTGGCCAGTGGGTGTCAGCAGTACTCTAAATATAGTGAAGCCGGGAAGGG 1560  
Qy 1561 CTCTCGTCAATGCTGCGTACAAATCTATGGCAGCTCTTGGCAACCATCTTTTA 1620  
Db 1561 CTCTCGTCAATGCTGCGTACAAATCTATGGCAGCTCTTGGCAACCATCTTTTA 1620  
Qy 1621 AACTCACCAGGCCCTTCCCTTTCATCCAAGACAACATCTCTTATCGCACCATTTGG 1680  
Db 1621 AACTCACCAGGCCCTTCCCTTTCATCCAAGACAACATCTCTTATCGCACCATTTGG 1680  
Qy 1681 CTCTGTCCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740  
Db 1681 CTCTGTCCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740  
Qy 1741 TTTAGGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGGCCCTTGATACCGAGTAC 1800  
Db 1741 TTTAGGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGGCCCTTGATACCGAGTAC 1800  
Qy 1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTTAAGTGGAGTTCCCGAGAGAGAAC 1860  
Db 1801 TTCTACGTTGACTTCAGGAGCTATGAATATGACCTTTAAGTGGAGTTCCCGAGTGAAGAC 1860  
Qy 1861 TTAGAGTTTGGGAAGGTCTGCGGTCTGGCGCTTTTCGGAGAGGTGATGAACCCACGCGC 1920  
Db 1861 TTAGAGTTTGGGAAGGTCTGCGGTCTGGCGCTTTTCGGAGAGGTGATGAACCCACGCGC 1920  
Qy 1921 TATGCAATAGTAAACGGGAGTCTCAATTCAGGTGGGGTGAAGATGCTAAAGAGAAA 1980  
Db 1921 TATGCAATAGTAAACGGGAGTCTCAATTCAGGTGGGGTGAAGATGCTAAAGAGAAA 1980  
Qy 1981 GCTGACAGTGTGAAGAAAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA 2040  
Db 1981 GCTGACAGTGTGAAGAAAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA 2040  
Qy 2041 CACCATGACAACATCGTGAATCTGCTGGGGGATGCACACTGTCAGGGCCAGTGTACTTG 2100  
Db 2041 CACCATGACAACATCGTGAATCTGCTGGGGGATGCACACTGTCAGGGCCAGTGTACTTG 2100

Qy 2101 ATTTTGAATATTGTTGCTATGCTGACCTCCTCAACTACCTTAAGAAAGTAAAGAGAGAAG 2160  
Db 2101 ATTTTGAATATTGTTGCTATGCTGACCTCCTCAACTACCTTAAGAAAGTAAAGAGAGAAG 2160  
Qy 2161 TTTACAGGACATGGACAGAGATTTTAAAGAAACATAAATTTCAAGTTTACCCTTACCTTC 2220  
Db 2161 TTTACAGGACATGGACAGAGATTTTAAAGAAACATAAATTTCAAGTTTACCCTTACCTTC 2220  
Qy 2221 CAGGACATTTCAATTTCCAGCATGCTGTTACAGAGAAAGTTCAAGTTTACACCCGCTTG 2280  
Db 2221 CAGGACATTTCAATTTCCAGCATGCTGTTACAGAGAAAGTTCAAGTTTACACCCGCTTG 2280  
Qy 2281 GATCAGCTCTCAGGTTCAATGGAATTTCAATTTCAATTTCAAGATGAGATTGAATATGAA 2340  
Db 2281 GATCAGCTCTCAGGTTCAATGGAATTTCAATTTCAATTTCAAGATGAGATTGAATATGAA 2340  
Qy 2341 AACCAGAAGAGGCTGCGCAGAAGAAGAGAGATTTTGAACGTGCTGAGCTTTGAAGAC 2400  
Db 2341 AACCAGAAGAGGCTGCGCAGAAGAAGAGAGAGATTTTGAACGTGCTGAGCTTTGAAGAC 2400  
Qy 2401 CTCCTTTGCTTTGCGTACCAAGTGGCCAAAGGATGGAATTTCTGGAGTTCAAGTCGTGT 2460  
Db 2401 CTCCTTTGCTTTGCGTACCAAGTGGCCAAAGGATGGAATTTCTGGAGTTCAAGTCGTGT 2460  
Qy 2461 GTCCACAGAGCTGCGCAGCCAGGAATGTTGTTGTCACCCACCGGAAGGTTGTTGAAGATC 2520  
Db 2461 GTCCACAGAGCTGCGCAGCCAGGAATGTTGTTGTCACCCACCGGAAGGTTGTTGAAGATC 2520  
Qy 2521 TGTGACTTTGGACTGCGCCGAGACATCTCTGAGCGACTTCCAGCTACGTCGTCAGGGCAAC 2580  
Db 2521 TGTGACTTTGGACTGCGCCGAGACATCTCTGAGCGACTTCCAGCTACGTCGTCAGGGCAAC 2580  
Qy 2581 GCACGGCTCCGGTGAAGTGGATGGCACCCGAGAGCTTATTGGAAGGGATCTACACAATC 2640  
Db 2581 GCACGGCTCCGGTGAAGGAGTGGCACCCGAGAGCTTATTGGAAGGGATCTACACAATC 2640  
Qy 2641 AAGAGTGAGCTCTGTCCTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTCAGC 2700  
Db 2641 AAGAGTGAGCTCTGTCCTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTCAGC 2700  
Qy 2701 CCTTACCCCTGGCATCTCTGTCGACGCTAACTTCTATAAAGTATTCAGAGTGGATTTAAA 2760  
Db 2701 CCTTACCCCTGGCATCTCTGTCGACGCTAACTTCTATAAAGTATTCAGAGTGGATTTAAA 2760  
Qy 2761 ATGGACGACCATTTCTATGCCACAGAGGATATCTTTGTAATGCAATCTCTGCTGGCT 2820  
Db 2761 ATGGACGACCATTTCTATGCCACAGAGGATATCTTTGTAATGCAATCTCTGCTGGCT 2820  
Qy 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTCTTTTAGGATGTCAGCTG 2880  
Db 2821 TTTGACTCAAGGAAGCGGCATCTTCCCAACCTGACTTCTTTTAGGATGTCAGCTG 2880  
Qy 2881 GCAGAGGCAAGAAAGCATGTATCAGAACATCCATCTACCAAAACAGGCGGCCCT 2940  
Db 2881 GCAGAGGCAAGAAAGCATGTATCAGAACATCCATCTACCAAAACAGGCGGCCCT 2940  
Qy 2941 CAGCAGAGAGGGGGCTCAGAGCCAGTCCACAGGCCAGGTCGAAGATTCACAGAGAA 3000  
Db 2941 CAGCAGAGAGGGGGCTCAGAGCCAGTCCACAGGCCAGGTCGAAGATTCACAGAGAA 3000  
Qy 3001 AGAAGTTAGCGAGGAGGCTTGACCCCGCCACCTTAGAGGGTGTAGACCCGAGAGCCA 3060  
Db 3001 AGAAGTTAGCGAGGAGGCTTGACCCCGCCACCTTAGAGGGTGTAGACCCGAGAGCCA 3060  
Qy 3061 AGATTAGCCTCGCTCTGAGGAAGCGCTTACAGCGCTTGGCTTGGCTGACTTTTCTCT 3120  
Db 3061 AGATTAGCCTCGCTCTGAGGAAGCGCTTACAGCGCTTGGCTTGGCTGACTTTTCTCT 3120  
Qy 3121 AGATGCTGTGCTGCTTACTTCCAAAGTGACTTCTATATAAATCAAACTCTCTCGCACAG 3180  
Db 3121 AGATGCTGTGCTGCTTACTTCCAAAGTGACTTCTATATAAATCAAACTCTCTCGCACAG 3180



QY 3181 GCGGGAGGCAATAAGAGACTTGTGGTGAGCGCGCTACCTGGGGGCTTTCCACG 3240  
DB 3181 GCGGGAGGCAATAAGAGACTTGTGGTGAGCGCGCTACCTGGGGGCTTTCCACG 3240  
QY 3241 AGCTTGAGGGGAAAGCCATGATCTGAAATATAGTATATCTTGTAAATAGCTGAACAA 3300  
DB 3241 AGCTTGAGGGGAAAGCCATGATCTGAAATATAGTATATCTTGTAAATAGCTGAACAA 3300  
QY 3301 ACCAAACCCGTTTGTGCTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTAA 3360  
DB 3301 ACCAAACCCGTTTGTGCTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTAA 3360  
QY 3361 AATACTATGTAACCTTTTTCATCTATTTAGTGATATATTTATGGATGAAATAAATCTTC 3420  
DB 3361 AATACTATGTAACCTTTTTCATCTATTTAGTGATATATTTATGGATGAAATAAATCTTC 3420  
QY 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453  
DB 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453

RESULT 9  
AAQ40914  
ID AAQ40914 standard; cDNA; 3453 BP.  
XX  
AC AAQ40914;  
DT 19-OCT-1993 (first entry)  
XX  
DE Murine flk-2 cDNA.  
XX  
KW Murine; receptor; protein; tyrosine kinase; ptk; flk-2; primitive;  
KW hematopoietic cell; mature; family; conserved; region;  
KW catalytic domain; c-kit; fetal liver kinase; flk; fetal; spleen;  
KW thymus; adult; brain; bone marrow; multipotential; CFU-Blast colony;  
KW hierarchy; transduction; T-lymphoid; lineage; ss.  
XX  
OS Mus musculus.  
XX  
FH Key  
XX CDS 31..3009  
FT Location/Qualifiers  
FT /\*tag= a  
FT sig\_peptide 31..111  
FT /\*tag= b  
FT mat\_peptide 112..3006  
FT /note= "Hydrophobic leader sequence"  
FT /\*tag= c  
XX  
PN W09310136-A.  
XX  
XX 27-MAY-1993.  
PD  
XX 16-NOV-1992; 92WO-US09893.  
PF  
XX 15-NOV-1991; 91US-0793065.  
PR  
XX (UYPR-) UNIV PRINCETON.  
PA  
XX Lemischka IR;  
PI  
XX WPI: 1993-182479/22.  
DR P-PSDB; AAR37502.  
DR  
XX Totipotent haematopoietic stem cell receptors, their ligands and  
PT DNA sequences - for treating anemia(s) and bone marrow damage  
PT due to e.g. cancer chemotherapy or radiotherapy  
XX  
PS Claim 6; Fig 1a; 127pp; English.  
XX  
XX This sequence encodes the murine receptor protein tyrosine kinase  
CC (PTK), flk-2. This nucleic acid is expressed in primitive hemato-  
CC poietic cells and not in mature hematopoietic cells. Members of  
CC this family of ptk's can be recognised by the conserved amino acid

CC regions in the catalytic domain. This family of ptk's also contains  
CC c-kit. These new receptors are termed fetal liver kinases (flk's)  
CC after the tissue in which they were discovered. flk-2 is also  
CC expressed in fetal spleen, fetal thymus, adult brain and adult  
CC bone marrow. flk-2 is expressed in individual multipotential CFU-  
CC Blast colonies capable of generating numerous multilineage colonies  
CC upon replating. It is likely therefore, that flk-2 is expressed in  
CC the entire primitive portion of the hematopoietic hierarchy. This is  
CC consistent with flk-2 being important in transducing putative self-  
CC renewal signals from the environment. flk-2 is the first receptor  
CC ptk known to be expressed in the T-lymphoid lineage.  
XX  
SQ Sequence 3453 BP; 947 A; 822 C; 875 G; 809 T; 0 other;  
Query Match 99.8%; Score 3446.6; DB 14; Length 3453;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3449; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 GCGGCTGGCTACCGCGGCTCCGGAGGCCATCGGGCGCTTGGCGCAGCGCAGCGACCGG 60  
DB 1 GCGGCTGGCTACCGCGGCTCCGGAGGCCATCGGGCGCTTGGCGCAGCGCAGCGACCGG 60  
QY 61 CGGCTGGCTGCTGCTGTTGTTTGTCAAGTAATGATTCTTTGAGACCCGTTTACAAACCAAGAC 120  
DB 61 CGGCTGGCTGCTGCTGTTGTTTGTCAAGTAATGATTCTTTGAGACCCGTTTACAAACCAAGAC 120  
QY 121 CTGCTGTGATCAAGTGTGTTTAAATCACTATGAGAACAAATGGCTCATAGCGGGAAG 180  
DB 121 CTGCTGTGATCAAGTGTGTTTAAATCACTATGAGAACAAATGGCTCATAGCGGGAAG 180  
QY 181 CCATCATCTACCGAATGTGGAGGATCCCGAAGAGCTCCAGTGTATCCCGAGCGGC 240  
DB 181 CCATCATCTACCGAATGTGGAGGATCCCGAAGAGCTCCAGTGTATCCCGAGCGGC 240  
QY 241 CAGAGTGAGGAGCGGTATATGAAGCGCCACCGTGGAGTGCCGAGTCTGGTCCATC 300  
DB 241 CAGAGTGAGGAGCGGTATATGAAGCGCCACCGTGGAGTGCCGAGTCTGGTCCATC 300  
QY 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGCGACCTTTCCTGCTCTGGGTCTTTAAGCAC 360  
DB 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGCGACCTTTCCTGCTCTGGGTCTTTAAGCAC 360  
QY 361 AGCTCCCTGGGTGCGCAGCGGCACCTTTGATTTACAAAACAGAGAAATCGTTCCATGCC 420  
DB 361 AGCTCCCTGGGTGCGCAGCGGCACCTTTGATTTACAAAACAGAGAAATCGTTCCATGCC 420  
QY 421 ATCTTGAAGTGCAGAGACCCAGGAGGAGATACCTACTCCATATTCAGAGCGAAGC 480  
DB 421 ATCTTGAAGTGCAGAGACCCAGGAGGAGATACCTACTCCATATTCAGAGCGAAGC 480  
QY 481 GCCAACTACACAGTACTGTTTCACAGTCAATGTAAAGAGATACAGCTGTATGCTAAGG 540  
DB 481 GCCAACTACACAGTACTGTTTCACAGTCAATGTAAAGAGATACAGCTGTATGCTAAGG 540  
QY 541 AGACCTTACTTTTAGGAAGATGGAACCCAGGATGCTGCTCTGATCTCCGAGGGGTGT 600  
DB 541 AGACCTTACTTTTAGGAAGATGGAACCCAGGATGCTGCTCTGATCTCCGAGGGGTGT 600  
QY 601 CCGGAGCCCACTGTGGAGTGGTCTCTGAGCTCCACAGGGAAGCTGTAAAGAGAA 660  
DB 601 CCGGAGCCCACTGTGGAGTGGTCTCTGAGCTCCACAGGGAAGCTGTAAAGAGAA 660  
QY 661 GGCCCTGCTCTTGTGAGAAAGAGGAGAAAGGTACTTTCATGTTGTTTCGGAACAGATC 720  
DB 661 GGCCCTGCTCTTGTGAGAAAGAGGAGAAAGGTACTTTCATGTTGTTTCGGAACAGATC 720  
QY 721 AGATGCTGTGCTAGAAATGCACTGGGCGCGCAATGCACCAAGCTGTTTCCACCATAGCTA 780  
DB 721 AGATGCTGTGCTAGAAATGCACTGGGCGCGCAATGCACCAAGCTGTTTCCACCATAGCTA 780  
QY 781 AACCAAGCTCCTCAGAGCACACTGCCCGAGTATTCTTGAAGTGGGGGACCCCTTGTGG 840  
DB 781 AACCAAGCTCCTCAGAGCACACTGCCCGAGTATTCTTGAAGTGGGGGACCCCTTGTGG 840

QY 841 ATCAGGTGTAGGCGCATCCATGTGAACCATCGATTTCGGCTCACCTGGAGCTGGAAGAC 900  
DB 841 ATCAGGTGTAGGCGCATCCATGTGAACCATCGATTTCGGCTCACCTGGAGCTGGAAGAC 900  
QY 901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACTCTCCACAAACAGACCAATG 960  
DB 901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACTCTCCACAAACAGACCAATG 960  
QY 961 ATTCCGATTCCTTTGGCCCTTTGTCCTTCCTGGGAAGAAACAGACCGGATATTACACC 1020  
DB 961 ATTCCGATTCCTTTGGCCCTTTGTCCTTCCTGGGAAGAAACAGACCGGATATTACACC 1020  
QY 1021 TGCTCTTCCTCAAAAGCACCCAGCAGTCACGCTGTGTGACCATCTAGAAAAGGGTTT 1080  
DB 1021 TGCTCTTCCTCAAAAGCACCCAGCAGTCACGCTGTGTGACCATCTAGAAAAGGGTTT 1080  
QY 1081 ATAAACGCTACCCAGCTCGCAAGAAGATGATAAATTTGACCCGTACGAAAAGTTCTGCTTC 1140  
DB 1081 ATAAACGCTACCCAGCTCGCAAGAAGATGATAAATTTGACCCGTACGAAAAGTTCTGCTTC 1140  
QY 1141 TCAGTCAGGTTTAAAGCGTACCCAGAAATCCGATGCGACGTGGATCTTCTCAAGCCCTCA 1200  
DB 1141 TCAGTCAGGTTTAAAGCGTACCCAGAAATCCGATGCGACGTGGATCTTCTCAAGCCCTCA 1200  
QY 1201 TTTCTCTGTGAACAGAGGCGCTGGAGATGGGTACACCATATCTAAATTTTGGCATCAT 1260  
DB 1201 TTTCTCTGTGAACAGAGGCGCTGGAGATGGGTACACCATATCTAAATTTTGGCATCAT 1260  
QY 1261 AAGAACAGCCAGAGAGTACATATCTATGCAAGAAATGATGACGCCAGTTCACCAAA 1320  
DB 1261 AAGAACAGCCAGAGAGTACATATCTATGCAAGAAATGATGACGCCAGTTCACCAAA 1320  
QY 1321 ATGTCAGGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380  
DB 1321 ATGTCAGGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380  
QY 1381 GCSTCTCTCTGATGGTACCCGCTACCCCTTGGACCTGGAAGAGTGTCCGAC 1440  
DB 1381 GCSTCTCTCTGATGGTACCCGCTACCCCTTGGACCTGGAAGAGTGTCCGAC 1440  
QY 1441 AAATCTCCCAATGTCACGAGGAAATCCAGAGGAGTTTGAATATAAGAAAGCTAACAGA 1500  
DB 1441 AAATCTCCCAATGTCACGAGGAAATCCAGAGGAGTTTGAATATAAGAAAGCTAACAGA 1500  
QY 1501 AAAGTGTGTGCCAGTGGGTGTGAGCAGTACTCTAAATATGAGTGAGGCGGGAAGGG 1560  
DB 1501 AAAGTGTGTGCCAGTGGGTGTGAGCAGTACTCTAAATATGAGTGAGGCGGGAAGGG 1560  
QY 1561 CTTCTGGTCAATGCTGTGCGTACAAATCTATGGCAGCTCTTGCGAAACCATCTTTTA 1620  
DB 1561 CTTCTGGTCAATGCTGTGCGTACAAATCTATGGCAGCTCTTGCGAAACCATCTTTTA 1620  
QY 1621 AACTCACCAGGCCCTTCCCTTTCATCAAGAACACATCTCCTCTATGCGACCATTTGG 1680  
DB 1621 AACTCACCAGGCCCTTCCCTTTCATCAAGAACACATCTCCTCTATGCGACCATTTGG 1680  
QY 1681 CTCGTCTCCCTTCATTTGTTGTTCTCATPTGTTGATGTGCCACAAATACAAAAGCAA 1740  
DB 1681 CTCGTCTCCCTTCATTTGTTGTTCTCATPTGTTGATGTGCCACAAATACAAAAGCAA 1740  
QY 1741 TTTAGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTTGGCCCTTGGATACAGATAC 1800  
DB 1741 TTTAGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTTGGCCCTTGGATACAGATAC 1800  
QY 1801 TTCTACGTTGACTTCAGGAGCTATCAATATGACCTTAAGTGGAGTTCCCGAGAGAGAC 1860  
DB 1801 TTCTACGTTGACTTCAGGAGCTATCAATATGACCTTAAGTGGAGTTCCCGAGAGAGAC 1860  
QY 1861 TTAGAGTTTGGAGGTCCTGGGCTGTGGCGCTTTTGGGAGGTTGATGAACCCAGGCC 1920  
DB 1861 TTAGAGTTTGGAGGTCCTGGGCTGTGGCGCTTTTGGGAGGTTGATGAACCCAGGCC 1920

QY 1921 TATGCGATTAGTAAACGGGAGTCTCAATTCAGGTGGGCTGAAGATGCTAAAGAGAAA 1980  
DB 1921 TATGCGATTAGTAAACGGGAGTCTCAATTCAGGTGGGCTGAAGATGCTAAAGAGAAA 1980  
QY 1981 GCTGACAGCTGTGAAAAAGAAAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040  
DB 1981 GCTGACAGCTGTGAAAAAGAAAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040  
QY 2041 CACCATGACACATCGTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTGTACTTTG 2100  
DB 2041 CACCATGACACATCGTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTGTACTTTG 2100  
QY 2101 ATTTTGAATATTTGTTGCTATGCTGACCTCTCAACTACTTAAGAGTAAAGAGAGAAG 2160  
DB 2101 ATTTTGAATATTTGTTGCTATGCTGACCTCTCAACTACTTAAGAGTAAAGAGAGAAG 2160  
QY 2161 TTTACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCACTTCTACCCCTACTTTC 2220  
DB 2161 TTTACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCACTTCTACCCCTACTTTC 2220  
QY 2221 CAGGACATTTCAAAATTCAGCATGCTGTTTCAAGAGAAAGTTTCACTTACCCCGCTTTG 2280  
DB 2221 CAGGACATTTCAAAATTCAGCATGCTGTTTCAAGAGAAAGTTTCACTTACCCCGCTTTG 2280  
QY 2281 GATCAGCTCTCAGGCTTCAATGGAATTCATTTCACTTGAAGATGAGATTTGAATATGAA 2340  
DB 2281 GATCAGCTCTCAGGCTTCAATGGAATTCATTTCACTTGAAGATGAGATTTGAATATGAA 2340  
QY 2341 AACCAAGAGAGGCTGGCAGAGAGAGAGAGAGATTTTGAAGCTGCTGAGTTTGAAGAC 2400  
DB 2341 AACCAAGAGAGGCTGGCAGAGAGAGAGAGAGATTTTGAAGCTGCTGAGTTTGAAGAC 2400  
QY 2401 CTCCTTTGCTTTGCTTACCAAGTGGCAAGGATGGAATTTCTGAGTTTCAAGTCTGCT 2460  
DB 2401 CTCCTTTGCTTTGCTTACCAAGTGGCAAGGATGGAATTTCTGAGTTTCAAGTCTGCT 2460  
QY 2461 GTCCACAGAGCTGCGCAGCAGAGATGTTGGTCAACCCAGGAGGTTGTTGAAGATC 2520  
DB 2461 GTCCACAGAGCTGCGCAGCAGAGATGTTGGTCAACCCAGGAGGTTGTTGAAGATC 2520  
QY 2521 TGTGACTTTGAGTGGCTGGCCGAGACATCTTCAAGCTTCCAGCTTACGCTGAGGGCAAC 2580  
DB 2521 TGTGACTTTGAGTGGCTGGCCGAGACATCTTCAAGCTTCCAGCTTACGCTGAGGGCAAC 2580  
QY 2581 GCACGGCTGCGGCTGAAGTGGATGGCACCAGAGAGCTTATTTGAAGGGATCTACAAATC 2640  
DB 2581 GCACGGCTGCGGCTGAAGTGGATGGCACCAGAGAGCTTATTTGAAGGGATCTACAAATC 2640  
QY 2641 AAGAGTGAGCTGCTGCTTACGCAATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700  
DB 2641 AAGAGTGAGCTGCTGCTTACGCAATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700  
QY 2701 CTTTACCCTGGCATCTGCTGACCGTAACTTCTTAAACTGATTTCAAGAGTGGATTTAA 2760  
DB 2701 CTTTACCCTGGCATCTGCTGACCGTAACTTCTTAAACTGATTTCAAGAGTGGATTTAA 2760  
QY 2761 ATGGAGCAGCCATTTATGCCACAGAGGATATCTTTGATGCAATCTCTCTGGCT 2820  
DB 2761 ATGGAGCAGCCATTTATGCCACAGAGGATATCTTTGATGCAATCTCTCTGGCT 2820  
QY 2821 TTTTACTCAAGGAAGCGGCATCTTCCCACTGACCTTCACTTTTATGATCTGAGCTG 2880  
DB 2821 TTTTACTCAAGGAAGCGGCATCTTCCCACTGACCTTCACTTTTATGATCTGAGCTG 2880  
QY 2881 GCAGAGGAGAGAGAGCATGATCAGAACATCCATCTTACCAAAACAGGCGGCCCT 2940  
DB 2881 GCAGAGGAGAGAGAGCATGATCAGAACATCCATCTTACCAAAACAGGCGGCCCT 2940  
QY 2941 CAGCAGAGAGGCGGCTCAGAGCCAGTGCACACAGCGCCAGGTGAAGATTCACAGAGAA 3000  
DB 2941 CAGCAGAGAGGCGGCTCAGAGCCAGTGCACACAGCGCCAGGTGAAGATTCACAGAGAA 3000  
QY 3001 AGAAGTTACGAGGAGGCGCTTGGACCCCGCCCTTAGCAGGCTGTAGACCGCAGAGCCA 3060

3001	AGAAGTTAGCGAGGAGCGCCTTGGACCCCGCACCCCTAGCAGGCTGTAGACCGCAGAGCCA	3060
3061	AGATTACGCTCGCCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCGCTGCAGCTTTTCTCT	3120
3061	AGATTACGCTCGCCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCGCTGCAGCTTTTCTCT	3120
3121	AGATGCTGTCTGCCATTACTCCAAGTGACTTCTATAAAATCAAACCTCTCCCTCGCACAG	3180
3121	AGATGCTGTCTGCCATTACTCCAAGTGACTTCTATAAAATCAAACCTCTCCCTCGCACAG	3180
3181	CGCGGAGAGCCAATAATGAGACTTGTGTGGTGAAGCCGCCCTACCCCTGGGGGCCCTTCCACG	3240
3181	CGCGGAGAGCCAATAATGAGACTTGTGTGGTGAAGCCGCCCTACCCCTGGGGGCCCTTCCACG	3240
3241	AGCTTGAGGGGAAGCCATCTACTCGAAATATAGTATATTCCTGTGTAATACGTGGAACAA	3300
3241	AGCTTGAGGGGAAGCCATCTACTCGAAATATAGTATATTCCTGTGTAATACGTGGAACAA	3300
3301	ACCAAAACCCGTTTTTTCCTAAGGGAAGCTAAATATGATTTTTTAAAAATCTATGTTTTAA	3360
3301	ACCAAAACCCGTTTTTTCCTAAGGGAAGCTAAATATGATTTTTTAAAAATCTATGTTTTAA	3360
3361	AATACTATGTAACCTTTTTCATCTATTTAGTGATATATTTTATGGATGGAATAAACCITTC	3420
3361	AATACTATGTAACCTTTTTCATCTATTTAGTGATATATTTTATGGATGGAATAAACCITTC	3420
3421	TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA	3453
3421	TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA	3453

RESULT 10

AA00801	standard; DNA; 3521 BP.
AA00801	
AA00801	
29-FEB-1996	(first entry)
Flk2/flt3	tyrosine kinase receptor gene.
Tyrosine kinase receptor; flk2; flt3; agonist; monoclonal antibody; haematopoiesis; hypoplasia; anaemia; thrombocytopenia; stem cell; ss.	
Mus sp.	
Key	Location/Qualifiers
CDS	82..3084
FT	/*tag= a
WO9527062-A1.	
12-OCT-1995.	
23-MAR-1995;	95WO-US03718.
04-APR-1994;	94US-0222299.
(GETH ) GENENTECH INC.	
Bennett BD, Broz SD, Matthews W, Zeigler FC;	
WPI: 1995-358636/46.	
P-PSDB; AAR81868.	
Agonist antibody which binds to flk2-flt3 tyrosine kinase receptor - enhances proliferation of haematopoietic stem cells, in the treatment of hypoplasia, anaemia, etc.	
Example 1; Page 34-37; 59pp; English.	
DNA (AA00801) coding for murine stem cell tyrosine kinase receptor flk2	

901 AAAGCCCTGGAGGAGGAGCTACTTTGATGAGTACCTACTCTACAAACAGGACCATG 960  
952 AAAGCCCTGGAGGAGGAGCTACTTTGATGAGTACCTACTCTACAAACAGGACCATG 1011  
961 ATTCGGATTCTTTGGCCCTTTGTGCTTCCGTTGGGAAGAACGACACCGGATATTACACC 1020  
1012 ATTCGGATTCTTTGGCCCTTTGTGCTTCCGTTGGGAAGAACGACACCGGATATTACACC 1071  
1021 TGCTCTTCTCTAAACACCCAGCCAGCTGAGCTGGTGACCATCTCTAGAAAAAGGTTT 1080  
1072 TGCTCTTCTCTAAAGACCCAGCCAGCTGAGCTGGTGACCATCTCTAGAAAAAGGTTT 1131  
1081 ATAAACGCTTACAGCTCGCAAGAGATGAAATTTGACCCGTAGCAAAAGTTCTGCTTC 1140  
1132 ATAAACGCTTACAGCTCGCAAGAGATGAAATTTGACCCGTAGCAAAAGTTCTGCTTC 1191  
1141 TCAGTCAGGTTTAAAGCGTTACCCAGATCCGATCCGATCGATGATCTTCTCAAGCCTCA 1200  
1192 TCAGTCAGGTTTAAAGCGTTTACCCAGATCCGATCCGATCGATGATCTTCTCAAGCCTCA 1251  
1201 TTTCTCTTGTGAACAGAGAGGCTTGGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
1252 TTTCTCTTGTGAACAGAGAGGCTTGGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1311  
1261 AAGAAACAGCCAGGAGAGTACATATTTCTATGCAGAAATGATGAGCCCGAGTTTCAACAAA 1320  
1312 AAGAAACAGCCAGGAGAGTACATATTTCTATGCAGAAATGATGAGCCCGAGTTTCAACAAA 1371  
1321 ATGTTACGCTGAATATAAGAAAGAACTCAAGTGTCTAGCAAAATGCTTCAAGCCAGCCAG 1380  
1372 ATGTTACGCTGAATATAAGAAAGAACTCAAGTGTCTAGCAAAATGCTTCAAGCCAGCCAG 1431  
1381 GGGTCTCTCTGATGGCTTACCCGCTACCCCTCTTGGACCTGGGAAGTGTTCGGAC 1440  
1432 GGGTCTCTCTGATGGCTTACCCGCTACCCCTCTTGGACCTGGGAAGTGTTCGGAC 1491  
1441 AAATCTCCCAATGGACGAGGAAATCCAGAGGAGTTTGGAAATAAAAGGCTACACAGA 1500  
1492 AAATCTCCCAATGGACGAGGAAATCCAGAGGAGTTTGGAAATAAAAGGCTACACAGA 1551  
1501 AAAGTGTGTTGGCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGAGGCGGGGAAGGG 1560  
1552 AAAGTGTGTTGGCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGAGGCGGGGAAGGG 1611  
1561 CTCTGTTGCAATGCTGTCGCTACATTTCTATGGGCACCTCTGCAAAACCATCTTTTA 1620  
1612 CTCTGTTGCAATGCTGTCGCTACATTTCTATGGGCACCTCTGCAAAACCATCTTTTA 1671  
1621 AACTCACCAGGCCCTTCCCTTTTATCCAGACACATCTCTCTATGCGACCATTTGGG 1680  
1672 AACTCACCAGGCCCTTCCCTTTTATCCAGACACATCTCTCTATGCGACCATTTGGG 1731  
1681 CTCTGTTCCCTTCAATGTTGTTCTCATTTGTTGATCTGGCAACAAATACAAAAAGCAA 1740  
1732 CTCTGTTCCCTTCAATGTTGTTCTCATTTGTTGATCTGGCACAATAACAAAAAGCAA 1791  
1741 TTTAGTACGAGATCAGCTGAGATGATCCAGGTGACCTGGCCCTGGATACAGGATAC 1800  
1792 TTTAGTACGAGATCAGCTGAGATGATCCAGGTGACCTGGCCCTGGATACAGGATAC 1851  
1801 TTTCTAGCTTGAATGAGTATGAAATGACCTTTAAGTGGGAGTTCCCGAGAGAGAAC 1860  
1852 TTTCTAGCTTGAATGAGTATGAAATGACCTTTAAGTGGGAGTTCCCGAGAGAGAAC 1911  
1861 TTTAGTATTTGGGAAGTCTCGGGGCTTGGCGCTTTCGGGAGGATGATGAACGCCAGGCC 1920  
1912 TTTAGTATTTGGGAAGTCTCGGGGCTTGGCGCTTTCGGGAGGATGATGAACGCCAGGCC 1971  
1921 TATGGCATTTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA 1980  
1972 TATGGCATTTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA 2031  
1981 GCTGACAGCTGTGAAAAAGAGCTCTCTATGTCGGACCTCAAAATGATGATGACCCACCTGGGA 2040

2032 GCTGACAGCTGTGAAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2091  
2041 CACCATGACAAATCTGTAATCTGCTGGGGCATGCACACTGTCAAGGGCCAGTGTACTTG 2100  
2092 CACCATGACAAATCTGTAATCTGCTGGGGCATGCACACTGTCAAGGGCCAGTGTACTTG 2151  
2101 ATTTTGAATATTTGCTTATGCTGACCTCTCAACTTACCTAAGAAAGTAAAGAGAGAAG 2160  
2152 ATTTTGAATATTTGCTTATGCTGACCTCTCAACTTACCTAAGAAAGTAAAGAGAGAAG 2211  
2161 TTTACAGGACATGACAGAGATTTTAAAGAAACATAAATTTACCTTCTTACCCTACTTTC 2220  
2212 TTTACAGGACATGACAGAGATTTTAAAGAAACATAAATTTACCTTCTTACCCTACTTTC 2271  
2221 CAGGCACATCAAAATTCACCATGCTGTTTACAGAGAAAGTTCAGTTTACACCCGCTTG 2280  
2272 CAGGCACATCAAAATTCACCATGCTGTTTACAGAGAAAGTTCAGTTTACACCCGCTTG 2331  
2281 GATCAGCTCTCAGGGTTCAATGGGAATTCATTTCAATTTCTGAAGATGAGATTGAATATCAA 2340  
2332 GATCAGCTCTCAGGGTTCAATGGGAATTCATTTCAATTTCTGAAGATGAGATTGAATATCAA 2391  
2341 AACCAGAAGAGGCTGCGAGAAGAGAGGAGGAAGATTTGAACGTGTGACGTTTGAAGAC 2400  
2392 AACCAGAAGAGGCTGCGAGAAGAGAGGAGGAAGATTTGAACGTGTGACGTTTGAAGAC 2451  
2401 CTCCTTTGCTTTCGTACCAAGTGGCCAAAGGATGGAATTTCTGGAGTTCAAGTCTGCT 2460  
2452 CTCCTTTGCTTTCGTACCAAGTGGCCAAAGGATGGAATTTCTGGAGTTCAAGTCTGCT 2511  
2461 GTCCACAGAGACCTGCGCAGCAGGAATGTGTTGGTCAACCCAGGGAAGTGGTGAAGATC 2520  
2512 GTCCACAGAGACCTGCGCAGCAGGAATGTGTTGGTCAACCCAGGGAAGTGGTGAAGATC 2571  
2521 TGTGACTTTGGACTGCGCCGAGACATCTCTGAGCGACTTCCAGCTACGTCGTCAGGGCAAC 2580  
2572 TGTGACTTTGGACTGCGCCGAGACATCTCTGAGCGACTTCCAGCTACGTCGTCAGGGCAAC 2631  
2581 GCACGCTCGCGTGAAGTGGATGGCCACCGAGAGCTTATTTGAAGGGATCTACACATC 2640  
2632 GCACGCTCGCGTGAAGTGGATGGCCACCGAGAGCTTATTTGAAGGGATCTACACATC 2691  
2641 AAGAGTGAAGTGTGCTTACGCGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700  
2692 AAGAGTGAAGTGTGCTTACGCGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2751  
2701 CCTTACCCTGGCATCTCTGTCGACGCTAACTTCTATAAAGTGTTCAGAGTGGATTTAAA 2760  
2752 CCTTACCCTGGCATCTCTGTCGACGCTAACTTCTATAAAGTGTTCAGAGTGGATTTAAA 2811  
2761 ATGGACAGCCATCTATGCCACAGAGGAGATATCTTTGTAATGCAATCTCTGCTGGCT 2820  
2812 ATGGACAGCCATCTATGCCACAGAGGAGATATCTTTGTAATGCAATCTCTGCTGGCT 2871  
2821 TTTGACTCAAGGAAGCGGCCATCTTCCCAACCTGACTTCTATTTTAGATGTGACGCTG 2880  
2872 TTTGACTCAAGGAAGCGGCCATCTTCCCAACCTGACTTCTATTTTAGATGTGACGCTG 2931  
2881 GCAGGCGAAGAAAGC-----ATGTTATCAGAACATCCATCC 2917  
2932 GCAGGCGAAGAAAGCGATGTATCAGAACATGGTGGCAACGTCGCCAACAATCCATCC 2991  
2918 ATCTACAAACAGGGGGCCCTCAGCAGAGAGCGGGCTCAGAGCC-CAGTCCGCCACAG 2976  
2992 ATCTACAAACAGGGGGCCCTCAGCAGAGAGCGGGCTCAGAGCGCCATCCGCCACAG 3051  
2977 CGCCAGGTGAAGATTCACAGAGAAAGATTTAGCGAGAGGCTTTGGACCCGCCACCT 3036  
3052 GCCCAGGTGAAGATTCACAGAGAAAGATTTAGCGAGAGGCTTTGGACCGGCCACCT 3111  
3037 AGCAGGCTGTAGACCGCAGAGCCAAAGATTAGCCCTCGCTCTGAGGAAGCGCCCTACAGCG 3096

Db 3112 AGCAGGCTGTAGACCGCAGAGCCAGGATTAGCCTCGCCTGTGAGGAAGCGCCCTACAGGC 3171  
Qy 3097 CGTTGCTTCGCTGGACTTTTCTCTAGATGCTGTCTGCCATTACTCCAAAGTGACTTCTAT 3156  
Db 3172 CGTTGCTTCGCTGGACTTTTCTCTAGATGCTGTCTGCCATTACTCCAAAGTGACTTCTAT 3231  
Qy 3157 AAAATCAAACTCTCCTCGCACAGCGGGAGAGCCAAATATGAGACTTGTGGTGAGGCC 3216  
Db 3232 AAAATCAAACTCTCCTCGCACAGCGGGAGAGCCAAATATGAGACTTGTGGTGAGGCC 3291  
Qy 3217 GCCTACCTGGGGGCTTT-----CCACGAGCTTGAGGGGAAGCCATGTATCTGAA 3268  
Db 3292 GCCTACCTGGGGGCTTTTCCAGGCCCTTCCAGGCCCTTGAGGGGAAGCCATGTATCTGAA 3351  
Qy 3269 ATATAGTATATCTTGTAAATACGTGAACAAACCAACCCGTTTTTTTGTAAAGGGAAG 3328  
Db 3352 ATATAGTATATCTTGTAAATACGTGAACAAACCAACCCGTTTTTTTGTAAAGGGAAG 3411  
Qy 3329 CTAATATGATTTTAAATATCTATGTTTTTAAATATCTATGTAATCTTTTTCATCTATTTA 3388  
Db 3412 CTAATATGATTTTAAATATCTATGTTTTTAAATATCTATGTAATCTTTTTCATCTATTTA 3471  
Qy 3389 GTGATATATTTATGGATGGAAATAAACTTTCTACTGTAAATAAAAAA 3438  
Db 3472 GTGATATATTTATGGATGGAAATAAACTTTCTACTGTAGAAAAA 3521

RESULT 11

AAQ81013 ID AAQ81013 standard; cDNA; 3501 BP.

XX AC AAQ81013;

XX DT 18-AUG-1995 (first entry)

XX DE Flk2 receptor protein-tyrosine-kinase cDNA.

XX KW Human Flk2; receptor protein-tyrosine-kinase; primitive

XX OS Homo sapiens.

FH Key

FT CDS 58..3039

FT /\*tag= a

FT /product= Flk2 receptor protein-tyrosine-kinase

FT sig\_peptide 58..138

FT mat\_peptide 139..3036

FT /\*tag= b

FT /\*tag= c

XX WO9500554-A.

XX 05-JAN-1995.

XX 17-JUN-1994; 94WO-US06944.

XX 18-JUN-1993; 93US-0080244.

XX 21-JUN-1993; 93US-0081508.

XX 23-NOV-1993; 93US-0157490.

XX (UYPR-) UNIV PRINCETON.

XX Lemischka IR;

XX WPI; 1995-052014/07..

XX P-PSDB; AAR67816.

XX Ligand for receptor protein tyrosine kinase - useful for the

XX stimulation of primitive haematopoietic stem cells causing

XX proliferation and/or differentiation

XX Disclosure; Fig 1b; 131pp; English.

XX PS

XX The sequence corresponds to a cDNA encoding a human Flk2 (fetal  
CC liver kinase) receptor protein-tyrosine-kinase. Flk2 is expressed  
CC in primitive hematopoietic cells but not in mature hematopoietic  
CC cells. The gene product is useful in isolation of receptor  
CC ligands, which have applications in diagnosis of bone marrow  
CC disorders and in stimulating proliferation and/or differentiation  
CC of primitive hematopoietic stem cells.  
XX SQ Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;

Query Match 64.5%; Score 2227.2; DB 16; Length 3501;  
Best Local Similarity 79.6%; Pred. No. 0;  
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

Qy 8 GGCTACCGCGCTCCGGAGGCGCATGCGGCGCTTGGCGAGCGAGCGCGCGGCTGC 67  
Db 35 GGGGACCCCGGGCTCCGGAGGCGCATGCGGCGCTTGGC---GGCGAGCGGGCACCGTGC 91  
Qy 68 TGCTGCTTGTGTTTGTGTCAGTAATGATTCCTTTGAGACCGTTTACAAACCAAGACCTGCCCTG 127  
Db 92 CGCTGCTGTTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCCCTG 151  
Qy 128 TGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGGGAAGCCATCAT 187  
Db 152 TGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCATCAGTGGGGAAGTGCATCAT 211  
Qy 188 CGTACCGAATGTCGCGAGGATCCCGAGAGACCTCCAGTGTACCCGCGCGCAGAGTG 247  
Db 212 CATATCCCATGTCATCAGAAATCCCGGAAGACCTCGGTGTGGTTGAGACCCAGAGCT 271  
Qy 248 AAGGACGCTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCTGGGTCTCATCACCTGC 307  
Db 272 CAGGACAGTGTACGAAGCTCGCGTGTGGAAGTGGATGTATCTGCTTCCATCACACTGC 331  
Qy 308 AAGTCCAGTCCGCCACCCAGCGGACCTTCTGCTCTGCGTCTTGGGTCTTTAAGCAGACTCCC 367  
Db 332 AAGTGTGTCGATGCCCGGAGAACATTTCTGCTCTGCGTCTTGGGTCTTTAAGCAGACTCCC 391  
Qy 368 TGGGCTGCCAGCGGCATTTGATTTACAAAACAGAGAAATCGTTTCCATGGCCATCTTGA 427  
Db 392 TGAATGCCAGCCACATTTGATTTACAAAACAGAGAGAGTGTGTTCCATGGTCTATTTGA 451  
Qy 428 ACCTGACAGAGACCCAGCGAGAGATACCTACTCCTCATATTACAGCGGAGCGGCGCAACT 487  
Db 452 AAATCACAGAAACCCCAAGCTGGAGATACCTACTTTTATTCAGAGTGAAGCTACCAATT 511  
Qy 488 ACACAGTACTGTTTCACAGTGAATGTAAGAGATACAGCTGTATGTCTGAAGAGACCTT 547  
Db 512 ACACAATATTTTACAGTGAATGTAAGAAATACCTGCTTTTACACATTAAAGAAGACCTT 571  
Qy 548 ACTTTAGGAAGATGGAACACAGGATGCACTGCTCTGCTCTCCGAGGCTGTTCCGAGGC 607  
Db 572 ACTTTAGAAAATGGAACACAGGAGCGCCCTGCTCTGCTCTGCTCTGAGAGGTTCCAGAGC 631  
Qy 608 CCACCTGTGGAGTGGGTGCTCTGCTGCTCCAGAGGAAAGCTGTAAAGAAGAGGCGGCTG 667  
Db 632 CGATCGTGAATGGGTGCTTTGCGATTCACAGGGGGAAGAGCTGTAAGAAGAAAGTCCAG 691  
Qy 668 CTGTTGTCAGAAAGGAGGAAAGGTAATCTCATGAGTTGTTGCGAACAGACATCAGATGCT 727  
Db 692 CTGTTGTTAAAAAGGAGGAAAGAGTCTTCATGATTTATTTGGGAGCGGACATAGGCTGT 751  
Qy 728 GTGCTAGAAATGCACCTGGCGCGGATGCACCAAGCTGTTCCACATAGATCTAAACCCAGG 787  
Db 752 GTGCCAGAAATGAACCTGGGAGGAAATGCACCAAGCTGTTTCAATAGATCTTAATCAAA 811  
Qy 788 CTCCTCAGACGACACTGCCCGCTTATTCCTGAAAGTGGGGGAAACCTTGTGGATCAGGT 847  
Db 812 CTCCTCAGACGACATTTGCCACATTTATTTTAAAGTAGGGGAAACCTTATGGATAGGT 871  
Qy 848 GTAAGGCCCATCCATGTGAACCATGATTCGGGCTCACCTGGGAGCTGGGAAGACAAAGCCC 907

Db 872 GCAAAGCTGTTCAATGTGAACCATGATTCGGGCTCACCTGGGAATTAGAAAAAAGCAC 931  
Qy TGGAGGAGGAGCTACTTTGAGATGAGTACCTACTTCCACAACAGGACCATGATTCGGA 967  
Db TCGAGGAGGCAACTACTTTGAGATGAGTACCTATTCAACAACAGAACTATGATACGGA 991  
Qy TTCCTTTGGCCTTTGCTGCTTCCTGGGGAAGGAACGACACCGGATATTACACCTGCTCT 1027  
Db TTCTGTTTGGCTTTGATCATCAGTGGCAAGGAACGACACCGGATATTACACCTGCTCT 1051  
Qy CCTCAAAACACCCAGCAGCTCAGCTGGGTGACCATCTCAAGAAAGGTTTATAAAGC 1087  
Db CTTCAAGCATCCCAAGTCAATCAGCTTTGGTTACCATCTGTAAGGAGGATTTATAAAG 1111  
Qy CTACAGCTCGCAAGAGATGATAATTTGACCCGTAGCAAAAGTTCTGCTTCTCAGTCA 1147  
Db CTACCAATTCAAGTGAAGATTATGAATGACCAATATGAAGATTTGTTTTCTGTCA 1171  
Qy GGTTTAAAGCGTACCACGAATCCGATGCGATGCGATGCTTCTCAAGCCTCATTTCCCT 1207  
Db GGTTTAAAGCCTACCCACAATTCAGATGTACGTGGACCTTCTCTCGAAATCATTTCCCT 1231  
Qy GTGAACAGAGGCTCGAGGATGGGTACAGCATATCTAAATTTTGGGATCATAGAACA 1267  
Db GTGAGCAAAAGGCTTGTATACGGATACAGCATATCCAAGTTTGGCAATCATNAGCAC 1291  
Qy AGCCAGGAGGTACATATCTATGAGAAATGATGACCCAGTTTACCAAAATGTTCA 1327  
Db AGCCAGGAGATATATATTCATGCAAGAAATGATGATGATGATGATGATGATGATGAT 1351  
Qy CGCTGAATATAGAAAGAACTCAAGTGTAGCAAAATGCTTACGCCAGCGGCTCT 1387  
Db CGCTGAATATAGAAAGAACTCAAGTGTAGCAAAATGCTTACGCCAGCGGCTCT 1411  
Qy GTTCTCTGATGGCTACCCGCTACCTCTTGGACCTGGAAGATGTTCCGGACAATCTC 1447  
Db GTTCTCGATGGATATACCATTTACCATCTTGGACCTGGAAGATGTTTCAGACAAGTCT 1471  
Qy CCAATTGACGAGGAAATCCAGAGGAGTTTGAATATAAAGGCTTAACAGAAAAGTCT 1507  
Db CCACTGGACAGAGATACAGAAAGGAGTCTGGAATAGAAGGCTTAACAGAAAAGTCT 1531  
Qy TTGCCAGTGGTGTGAGCAGTACTTAAATATAGTGTAGCGCGGGAAGGCTTCTGG 1567  
Db TTGGACAGTGGTGTGAGCAGTACTTAAACATGAGTGAAGCAATAAAGGCTTCTGG 1591  
Qy TCAATGCTGCGTACAATTTCTATGGCAGCTCTTGGCAACCATCTTTTAAACTCAT 1627  
Db TCAAGTGTGTGCATACAAATTTCCCTTTGGCAGATCTTGTGAGACGATCTCTTTAA 1651  
Qy CAGGCCCTTCCCTTTTCAATCAAGACAACATCTCTCTATGCGACCATTTGGGCTCTG 1687  
Db CAGGCCCTTCCCTTTTCAATCAAGACAACATCTCTCTATGCGACCATTTGGGCTCTG 1711  
Qy TCCCTTCTATTTGTTCTCATTTGTTGATCGCACAAATACAAAAGCAATTTAGGT 1747  
Db TCCCTTCTATTTGTTCTCATTTGTTGATCGCACAAATACAAAAGCAATTTAGGT 1771  
Qy ACGAGATGAGTGCAGATGATCCAGTGTAGTGGCCCCCTGGATTAACGATCTCTACG 1807  
Db ATGAAGCAGCTACAGATGGTACAGTGTAGGCTGACCGGCTCTCATGATGAGTCTCTAC 1831  
Qy TTGACTTTCAGGACTATGAATGACCTTAAGTGGGAGTTCCGAGAGAGACTTAGAGT 1867  
Db TTGATTTTCAGAGAATATGAATGATGATCAATGGGAGTTTCCAGAGAAATTTAGAGT 1891  
Qy TTGGGAAGTCTCTGGGCTGTGGCTTTCGGGAGGTTGATGAACCCAGCGCTATGGCA 1927  
Db TTGGGAAGTCTCTAGGATCAGGCTGCTTTTGGAAAGTGTATGAACCAACAGCTTATGGA 1951  
Qy TTGTTAAACGGGAGTCTCAATTTCAAGTGGCGGTGAAGATGCTAAAAGAGAAAGTGA 1987  
Db TTAGCAAAACAGGAGTCTCAATTCACAGTTGCGGTCAAAATGCTGAAAGAAAGCAGACA 2011

Qy 1988 GCTGTAAAAAAGAGCTCTATGTCGGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047  
Db 2012 GCTGTAAAGAGAGGACCTACTGTCAAGAACTCAAGATGATGACCCAGCTGGGAAGC 2071  
Qy 2048 ACAACATCTGATCTGCTGGGGCATGACACTGTCAAGGCCAGTGCTACTGATTTTGG 2107  
Db 2072 AGAATATTGTGAACCTGCTGGGGCGTGCACACTGTCAAGGACCAATTTACTGATTTT 2131  
Qy 2108 ATATTGTTGCTATGCTGACCTCTCAACTTACCTAAGAACTAAAGAGAGAAAGTTTCA 2167  
Db 2132 ATACTGTTGCTATGCTGATCTTCTCACTATCTAAGAACTAAAGAGAAATTTTCA 2191  
Qy 2168 GGACATGGACAGAGATTTTAAAGAACTAATTTTCACTTCTTACCTTACTTTCCAGG 2227  
Db 2192 GGACTTGGACAGAGATTTTCAAGAACTAATTTTCACTTCTTACCTTACTTTTCA 2251  
Qy 2228 ATTCAAATTCAGATGCTGCTTACAGAGAGTTCAGTTTACACCCGCCCTTGGATCAG 2287  
Db 2252 ATCCAAATTCAGATGCTGCTTCAAGAACTTCAAGAACTTCAAGAACTTCAAGAA 2311  
Qy 2288 TCTCAGGCTCAATGGGAATTTCAATTCATTTCTCAAGATGAGATTGAATATGA 2347  
Db 2312 TCTCAGGCTTCAATGGGAATTTCAATTCATTTCTCAAGATGAGATTGAATATGA 2371  
Qy 2348 AGAGGCTGCGAAGAAAGAGGAGGAGATTGAACGCTGCTGAGCTTTGAAGACCTCT 2407  
Db 2372 AAAGGCTG-----GAAGAGAGGAGGACTTGAATGCTTACATTTGAAGATCTCT 2425  
Qy 2408 GCTTTCGTCACCAAGTGGCCAAAGCATGGAATTCCTGGAGTTCAAGTCTGCTGTGTC 2467  
Db 2426 GCTTTCGTCATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCTGTGTC 2485  
Qy 2468 GAGACCTGCGACCCAGCAATGCTTGGTTCACCCAGCGGAGGTTGGTCAAGATCTG 2527  
Db 2486 GAGACCTGCGCCCGCAGGAACGCTGCTTGTACCCAGCGGAGGTTGGTGAAGATG 2545  
Qy 2528 TTGGACTGCGCCGAGACATCTTCAAGCAGCTTCCAGCTACGTCGTCAGGGGCAAG 2587  
Db 2546 TTGGATTTGCTCGAGATATCATGAGTGTATCCAACTATGTTGTCAAGGGCAAT 2605  
Qy 2588 TGCCGCTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACAAATCAAGAG 2647  
Db 2606 TGCTGTAAATGAGTGGTGGCCCGGAAAGCTGTTTGAAGGATCTACACCATTAAGAG 2665  
Qy 2648 AGCTGTGCTTACGSCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAACCTTACC 2707  
Db 2666 ATGCTGTGTCATATGGAATATTACTGTGGAAATCTTCTCACTGGTGTGAATCTTACC 2725  
Qy 2708 CTGGCATTTCTGTCGACGCTAACTTCTATAAAGTATTCAGAGTGGATTTAAATG 2767  
Db 2726 CTGGCATTTCTGTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 2785  
Qy 2768 AGCCATTTCTATGCCAGAGGAGATATCTTTGTAATGCAATCTCTGCTGGCTTTT 2827  
Db 2786 AGCCATTTCTATGCTACAGAAAGAAATATACATTTAATGCAATCTCTGCTGGCT 2845  
Qy 2828 CAAGAGGCGGCATCTTCCCAACCTGACITTCATTTTATGATGTCAGCTGGCAGG 2887  
Db 2846 CAAGAGAGGCGCATCTTCCCTTAATTTGACTTCGTTTATGATGTCAGCTGGCAG 2905  
Qy 2888 CAGAGAGAG-----CATGTATCAGAAACATCCATCATCTTACC 2924  
Db 2906 CAGAGAGAGGATGATCAGAAATGTCGATGGCGCTGTTTCGGAATGTCTTCACACCT 2965  
Qy 2925 AAAACAGCGCGCCTCAGCAGAGAGCGGCTCAGAGC--CCAGTGGCCACAGCGCCAG 2983  
Db 2966 AAAACAGCGCCTTTCAGCAGAGAGATGATTTGGGGTACTCTCTCCCGAGCTCAG 3025  
Qy 2984 T--GAAGATTCAGAGAAAGAAAGTTAGCGAGGAGGCTTTGGACCCCGCC-----AC 3038  
Db 3026 TCGAAGATTCGTAGAGGAACAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCT 3085





Qy	788	CTCCTCAGAGCAGACTGCCCCAGTATTCTCTGAAAGTGGGGGAACCCCTTGTGGATCAGGT	847
Db	812	CTCCTCAGAGCAGCATTCGCCAATATTATTTCTTAAAGTAGGGGAACCCCTTATGGATAGGT	871
Qy	848	GTAAGGCCATCCATGTGAACCATGTGGATTTCGGGCTCACCTCGGAGCTGGGAAGACAAAGCCC	907
Db	872	GCAAAGCTGTTCATGTGAACCATGGATTTCGGGCTCACCTCGGAATTAGAAGAACAAAGCAC	931
Qy	908	TGGAGGAGGCAGCTACTTTTGAGATGAGTAGTACCTACTCCACAAACAGGACCATGATTCGGGA	967
Db	932	TCGAGGAGGGCAACTACTTTTGAGATGAGTAGTACCTATTCCAACAAACAGAACTATGATACGGGA	991
Qy	968	TTCTCTTGGCCCTTTTGTCTCTTTCGGTGGGAAGGAACGACACCGGATATTACACCTGCTCTT	1027
Db	992	TTCTGTTTTGGTTTTTGTATCATGAGTGGCAAGAACGACACCGGATACACTGTGTCTCT	1051
Qy	1028	CCTCAAAAGCACCCAGCCAGCTCAGCGTTTGTGTGACCATTCCTAGAAAAAGGTTTATAAAGC	1087
Db	1052	CTTCAAAAGCATCCAGTCAATCAGCTTTGGTTACCATCTGTAGGAAGGGATTTATAAATG	1111
Qy	1088	CTACCGACTCCGAAGAAGATGATGAATTGACCCGTACGAAAAAGTTCTGCTCTCTCAAGCTCA	1147
Db	1112	CTACCAATTCAAGTGAAGATTATGAATTCACCAATATGAAGATTTTGTGTTCTCTGTCGA	1171
Qy	1148	GGTTTAAAGCGTACCCAGGAATCGGATGACGTGGATCTTCTCTCAAGCCCTCAATTTCCCTT	1207
Db	1172	GGTTTAAAGCGTACCCCAAAATCAGATGTACGTGGACCTTCTCTCGAAAAATCATTTCCCTT	1231
Qy	1208	GTAAGCAGAGAGCGCTCGAGGATGGGTACAGCATATCTAAATTTTGGCGATCATAGAAGCA	1267
Db	1232	GTGAGCAAAAGGCTCTTGATAAGCGGTACAGCATATCCAAAGTTTGGCAATCATAGCAACC	1291
Qy	1268	AGCCAGGAGAGTACATATCTATTCAGAAAAATGATGACGCCACGTTCACCAAAATGTTCA	1327
Db	1292	AGCCAGGAGAAATATATATTCATGCAAAAAATGATGATGCCAATTTTACCAAAATGTTCA	1351
Qy	1328	CGCTGATATTAAGAAAGAAACCTCAAGTGTCTAGCAATGCTCAGCCAGCAGCGGCTCCT	1387
Db	1352	CGCTGAAATTAAGAAGGAACCTCAAGTGTCTCGCAGAAGCATCGGAAGTCAGGCGCTCT	1411
Qy	1388	GTTCCCTGTGATGGCTACCCGCTACCCCTTTTGGACCTGGAAGAAGTGTTCGGACAATCTC	1447
Db	1412	GTTTCTCGGATGGATACCCATTACCATCTTGGACCTGGAGAAGTGTTCAGACAGTCTC	1471
Qy	1448	CCAATTGACGGAGGAATCCCAAGAGGATTTGGATAAAAAAGGCTTAACAGAAAAAGTGT	1507
Db	1472	CCAACCTCAGAGAAGATCACAGAAGAGTCTGGAATACAAAAGGCTTAACAGAAAAAGTGT	1531
Qy	1508	TTGSCCAGTGGGTCTCGACGAGTACTCTAAATATGAGTGAAGCCGGGAAGGCTCTCTGG	1567
Db	1532	TTGGACAGTGGGTCTCGACGAGTACTCTAAACATGAGTGAAGCCATAAAAGGGTTCCTCGG	1591
Qy	1568	TCAATGCTCTGGGTACAATTTCTATGGGCACGCTTTCGGAAGCCATCTTTTTTAAACTCAC	1627
Db	1592	TCAAGTCTCTGCATACAATTTCCCTTTGGCACAATCTTGTGAGACGATCCTTTTAAACTCTC	1651
Qy	1628	CAGGCCCTTCCCTTTCATCCAGACAAACATCTCCTTCTATGAGCAATTTGGGCTCTGTC	1687
Db	1652	CAGGCCCTTCCCTTTCATCCAGACAAACATCTCATTTATGCAACAATTTGGTGTGTTGTC	1711
Qy	1688	TCCCTTCATTGTGTCTCATTTGTGTGATCTGCCACAATACAAAAAGCAATTTAGGT	1747
Db	1712	TCCTCTTCATTGTCGTTTTAACCTTGCATAATTTGTTCACAAAGTACAAAAAGCAATTTAGGT	1771
Qy	1748	ACGAGTCAAGTCAAGTATGCAATATGACCTTAAGTGGGAGTTCGCCAGAGAGAACTTAGAGT	1807
Db	1772	ATGAAAGCCAGCTACAGATGGTACAGTCAAGCTACCGGCTCCTCTCAGATAATGAGTACTTCTACG	1831
Qy	1808	TTGACTTCAAGGACTATGAATATGACCTTAAGTGGGAGTTCGCCAGAGAGAACTTAGAGT	1867
Db	1832	TTGATTTTCAGAGAAATGAATATGATCTCAAAATGGGAGTTTCCAAGAGAAAAATTTAGAGT	1891

Qy	1868	TTGGGAAGGTCCTGGGGCTCTGGCGCTTTTCGGGAGGGTGATGAACGCCACCGCCCTATGGCA	1927
Db	1892	TTGGGAAGGTAAGGATAGGATAGGCTCTTTTGGAAAAGTGATGAACGCCAACAGCTTATGGAA	1951
Qy	1928	TTAGCTAAACGGGAGTCTCAATTCAGGTGCGGTGAAGATGCTAAAAGAGAAAGCTGACA	1987
Db	1952	TTAGCAAAACAGAGTCTCAATCCAGGTTGCCGTCAAATGCTGAAAGAAAAGCAGACA	2011
Qy	1988	GCTGTGAAAAGAGCTTCTCATGTCGGAGCTCAAAATGATGACCCACTGGGACACCATG	2047
Db	2012	GCTCTGAAAGAGGCACTCATGTCAAGACTCAAGATGATGACCCAGCTGGGAAGCCACG	2071
Qy	2048	ACAACTCATGTGAATCTGCTGGGGCATGACACTGTGACGGGCCAGTGACTCTGATTTTGTG	2107
Db	2072	AGAATATGTGAACCTCTCTGGGGCGTGCACACTGTTCAGGACCAATTTACTTGATTTTGTG	2131
Qy	2108	AATATTCTGCTATGTTGACCTCCTCAACTACCTAAGAAGTAAAGAGAGAAAGTTTTCACA	2167
Db	2132	AATACTGTTGCTATGTTGATCTTCTCAACTATCTTAAGAAGTAAAGAGAAAAATTTTCACA	2191
Qy	2168	GGACATGGACAGAGATTTTAAAGAAACATAAATTTCAAGTTCTTACCTTACTTTCCAGGCAC	2227
Db	2192	GGACTTGGACAGAGATTTTCAAGGAACACAATTTCAAGTTTACCACACTTTCCAATCAC	2251
Qy	2228	ATTCAAATTCAGATCGCTTGGTTTCAGGAGAAGTTTCAGTTTACACCCGCCCTTGGATCAGC	2287
Db	2252	ATCCAAATTCAGCATCGCTGGTTTCAAGAGAAAGTTTCAGATACACCCGGACTCGGATCAA	2311
Qy	2288	TCTCAGGTTCAATGGGAATTCATTTCAATTCATTTCTGNAAGATGAGATGAATATGAANAACGA	2347
Db	2312	TCTCAGGGCTTCATTGGGAATTCATTTCCACTCTCAAGATGAAATTTGAATATGAAANCCAAA	2371
Qy	2348	AGAGCTGGCAGAAGAAGAGGAGGAAGATTTGAACGTGCTGAGCTTTTGAAGACCTCCTTT	2407
Db	2372	AAAGCTG-----GAAGAGAGAGGAGACTTGAATGTGCTTACATTTGAAGATCTCTTT	2425
Qy	2408	GCTTTGCGTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAAGTCTGTGTCCACA	2467
Db	2426	GCTTTGTCATATCAAGTTTGCCAAAGAAATGAAATTTCTGGAATTTAAGTCTGTGTTTCA	2485
Qy	2468	GAGACCTGGCAGCAGCAATGTGTTGTCACCACCGGAAGGTGGTGAAGATCTGTGACT	2527
Db	2486	GAGACCTGGCGCGCCAGGAAGCTGCTGTCAACCACCGGAAGGTGGTGAAGATATGTGACT	2545
Qy	2528	TTGACATGGCCGCGAGACATCCTGAGGCATCCAGCTACGTACGTGAGGGCAACGCACGCG	2587
Db	2546	TTGATTTGGCTCGAGATATCATGATGATTCACAACTATGTGTGTCAGGGGCAATGCCCGTC	2605
Qy	2588	TGCCGTTGAATGGATGGCACCOCGAGAGCTTATTTGAAGGGATCTACACAATCAAGAGTG	2647
Db	2606	TGCCGTGAAAATGGATGGCCCCGCGAAAGCCTGTTTGAAGGCATCTACACCATTAAAGAGTG	2665
Qy	2648	ACGTCTGGTCTACGGCATCCTTCTCTGGGAGATATTTTTCACGTGGGTGTGAACCCCTTACC	2707
Db	2666	ATGTCTGGTCATATGGAATATTAATCTGGGAAATCTTCTCACTTGGTGTGAAATCCCTTACC	2725
Qy	2708	CTGGCATTCCTGTGCGACCTAACTTCTATAAAGTATTCAGATGGATTTTAAATGGAGC	2767
Db	2726	CTGGCATTCGGTGTGATGCTTAATCTCAAACTGATTCAAATGGATTTAAATGGATC	2785
Qy	2768	AGCCATTTCTATGCCACAGAAGGATATPACTTTGTATGCAATCCTGCTGGGCTTTTGTACT	2827
Db	2786	AGCCATTTTATGCTACAGAAGAAATATACATTTAATGCAATCCTGCTGGGCTTTTGTACT	2845
Qy	2828	CAAGGAAGCGGCATCCTTCCCACACCTGACTTCATTTTATAGATGTCACCTGGCAGAGG	2887
Db	2846	CAAGGAACGGCCATCCTCCCTAAATTTGACTTCGTTTATAGATGTCACCTGGCAGATG	2905
Qy	2888	CAGAAGAAG-----CATGTATCAGAACATCCCATCTCATCTTACC	2924
Db	2906	CAGAAGAAGGATGTATCAGAAATGTGATGGCGTGTTCGGAATGTCTTCACACACTTACC	2965
Qy	2925	AAAACAGGCGGCCCTTCAGCAGAGAGCGGGCTCAGAGC--CCAGTCCGCCACAGCGCCAGG	2983

```
Db 2966 AAAACAGCGACCTTTCAGAGAGAGATGGATTTGGGGCTACTCTCTCCGAGCGCTCAGG 3025
Qy 2984 T-GAAGATTACAGAGAAGAAAGTTAGCAGAGAGCGCTTGGACCCCGCC-----ACCCTAG 3038
Db 3026 TCGAAGATTTCGTAAGGAAACAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAA 3085
Qy 3039 CAGCGTGTAGACCGCAGAGCAAGATTAGCTCGGCTCT--GAGGAAGCGCCCTACAGCG 3096
Db 3086 CAGGCTGTAGATTACCAAAACAAGATTAAATTCATCACTAAAGAAAATCTATTATCAAC 3145
Qy 3097 CGTTCCTCGCTGGACTTTTCTCTAGATCGTCTGCCATTACTC-----CAAAGTGA 3149
Db 3146 TGCTGCTTACCAGACTTTTCTAGAGCGCTCTGCGCTTACTCTTCTTTTCAAGGGA 3205
Qy 3150 CTTCTATAAATCAAACTCTCTCGCACAGCGGGGAGAGCAATAATGAGACTTGTGTGG 3209
Db 3206 CTTTGTGAAAATCAATCATCTGTCTCAAGGCGAGGAGCTGATAATGAACTTTATTG 3265
Qy 3210 TGAGCCGCGCTACCTCGGGGCTTTCCACGAGCTTGAGGGAAAGCCATGTATCTGAAA 3269
Db 3266 GAGCATTTGATCGATCCAGCGCTTTCAGCGCGCTTGAGTGAATGTGTACTGTGAAG 3325
Qy 3270 TATAGTATATTCTTGTAAATACGTGAAGAAACAAACCAACCCGTTTTTTTCTTAAGGAAAGC 3329
Db 3326 TACAGTATATTCTTGTAAATACATAAAACAAA-----AGCATTTTCTTAAGGAGAACG 3378
Qy 3330 TAAATATGATTTTAAATATCTATGTTTTTAAATACATGTAACATTTTTCATCTATTAG 3389
Db 3379 TAAATATGATTTT--AAGTCTATGTTTTTAAATATATATGTAATTTTTCAGCTATTAG 3436
Qy 3390 TGATATATTTTATGATGGAATAAATCTTCTACTGTAATAAATAAATAAATAAATAAATAA 3449
Db 3437 TGATATATTTTATGCGTGGATATAAATTTCTACTACAGAAAAAATAAATAAATAAATAA 3496
Qy 3450 AAAA 3453
Db 3497 AAAA 3500

RESULT 13
AAT38734
ID AAT38734 standard; cdna; 3501 BP.
AC AAT38734;
XX
DT 11-DEC-1996 (first entry)
DE Murine foetal liver kinase 2 cdna.
XX
KW Murine; foetal liver kinase 2; flk-2; protein tyrosine kinase;
KW monoclonal; antibody; extracellular domain; receptor assay;
KW haematopoietic stem cell; ligand; stimulation; proliferation;
KW differentiation; treatment; anaemia; bone marrow damage;
KW cancer chemotherapy; radiation; ds.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 58..3039
FT FT /*tag= a
FT FT sig_peptide 58..138
FT FT /*tag= b
FT FT mat_peptide 139..3036
FT FT /*tag= c
XX
XX US5548065-A.
XX
XX 20-AUG-1996.
XX
XX 02-APR-1991; 91US-0679666.
XX
XX 19-NOV-1992; 92US-0977451.
```

```
PR 02-APR-1991; 91US-0679666.
PR 28-JUN-1991; 91US-0728913.
PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
PR 26-JUN-1992; 92US-0908397.
PR 12-NOV-1992; 92US-0975049.
PR 30-APR-1993; 93US-0055269.
PR 31-OCT-1994; 94US-0252517.
XX (UYPR-) UNIV PRINCETON.
XX
XX Lemischka IR;
XX
XX WPI; 1996-392678/39.
XX P-PSDB; AAR97419.
XX
XX Anti-foetal liver kinase 2 (flk-2) antibodies - useful in assays,
XX for isolating haematopoietic stem cells expressing receptor and for
XX obtaining ligands
XX
XX Disclosure; Columns 39-48; 50pp; English.
XX
XX The present sequence encodes murine foetal liver kinase 2 (flk-2),
XX a protein tyrosine kinase. Isolated antibodies, pref. monoclonal,
XX raised against the extracellular portion of flk-2 can be used to
XX assay for flk receptors on the surface of primitive haematopoietic
XX stem cells, and to isolate positive cells. The antibodies can also
XX be used as, or to obtain ligands, which stimulate the proliferation
XX and/or differentiation of stem cells. The ligands can be used, e.g.
XX for treating anaemia, or bone marrow damage resulting from cancer
XX chemotherapy, or radiation.
XX
XX Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;
XX
Query Match 64.5%; Score 2227.2; DB 17; Length 3501;
Best Local Similarity 79.6%; Pred. NO. 0;
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;
Qy 8 GGCTACCGCGCGCTCCGAGGCGCATCGGCGCTTGGCGCAGCGAGCGCGCGCTGC 67
Db 35 GGGGACCCCGCGCTCCGAGGCGCATCGGCGCTTGGC---GCGGAGCGCGGACCGTGC 91
Qy 68 TGCTGCTGCTGTTGTTTGTGAGTAAATGATTTCTGAGACCGTTTACAAACCAAGACCTGCTG 127
Db 92 CGCTGCTGCTGTTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCTG 151
Qy 128 TGATCAAGTGTGTTTAAATCAAGTCAATGAGACATGCTCATCAGCGGGAAGCCATCAT 187
Db 152 TGATCAAGTGTGTTTAAATCAATCAAGAACAAATGATTCATCAGTGGGGAAGTCAATCAT 211
Qy 188 CGTACCGAATGGTGGCAGGATCCCGAGAGACCTCCAGTCTAGCCCGGCGGAGAGTG 247
Db 212 CATATCCCATGGTATCAGAAATCCCGGAGAGACCTCGGCTGCTGAGACCCAGAGCT 271
Qy 248 AAGGGACGCTATATGAAGCGCGCACCGTGGAGTGGCCGAGTCTGGGTGCATCACCTGC 307
Db 272 CAGGGACAGTGTACGAAGCTGCCGCTGTGGAAGTGGATGATCTGCTTCCATCACACTGC 331
Qy 308 AAGTGCAGCTCGCCACCCCGAGGAGCTTTTCCTGCTCTGCTGCTTTTAAACACAGCTCCC 367
Db 332 AAGTGTGCTGATGATGCCCGAGGGAACATTTTCTGCTCTGCTGCTTTTAAACAGCTCCC 391
Qy 368 TGGGCTGCCAGCGCGCACTTTGATTTTACAAACAGAGGAATCGTTTCCATGGCCATTTGA 427
Db 392 TGAATTGCCAGCCACATTTTGATTTTACAAACAGAGAGTGTGTTCCATGGTCAATTTGA 451
Qy 428 AGTGCAGAGACCCCGAGGAGAGATACCTTACTTCCATATTCAGAGCGAAGCCGCAACT 487
Db 452 AAATGACAGAACCCCAAGCTGGAGATACCTACTTTTATTCAGAGTGAAGCTACCAAT 511
Qy 488 ACACAGTACTGTTTCAGAGTGAATGTAAAGATACAGAGCTGTATGTCTTAAAGAGACCTT 547
Db 512 ACACAAATATGTTTACAGTCAAGTAAAGAAATACCTGCTTTACACATATTAAAGAGACCTT 571
```

QY 548 ACTTTAGGAAGATGGAAGAACCCAGGATGCACCTGCTGCTGCTCGAGGGTGTTCGGGAGC 607  
DB 572 ACTTTAGAAAATGGAAGAACCCAGGACGCCCTGGTCTGCTCATATCTGAGAGCGTTCCAGAGC 631  
QY 608 CCACTGTGGAGTGGGTCTCTGACGCTCCACAGGGAAGCTGTAAGAGAGAGAGCCCTG 667  
DB 632 CGATCGTGAAGTGGGTCTTGGGATTACAGGGGGGAAGCTGTAAGAGAGAGAGTCCAG 691  
QY 668 CTGTTGTCAAGAGGAGGAAGAGTACTTTCATGAGTTGTTCCGGAACAGACATCAGATGCT 727  
DB 692 CTGTTGTAAAGAGGAGGAAGAGTGTCTCATGAATATTGAGGACGGACATAAGTGTCT 751  
QY 728 GTGCTAGAAATGCACTGGGCGCGAATGCACCAAGCTGTTCCACCATAGATCTAAACAGG 787  
DB 752 GTGCCAGAAATGAAGTGGGAGGGAATGCACAGGCTGTTCCACATAGATCTAAATCAA 811  
QY 788 CTCTCAGACACACTGCCCCAGTTATTCTGAAAGTGGGGAACCCCTGTGGATCAGGT 847  
DB 812 CTCTCAGACACACTGCCACAATTATTCTTAAAGTAGGGGAACCCCTTATGGATAAGT 871  
QY 848 GTAAGGCCATCCATGTGAACCATGGATTCGGGCTACCTGGGAGCTGGAAGACAAAGCC 907  
DB 872 GCAAGCTGTTCATGTGAACCATGGATTCGGGCTACCTGGGAATTAGAAAACAAAGCAC 931  
QY 908 TGGAGGAGGCGAGTACTTTCAGATGAGTACCTACTCCACAAACAGGACCATGATTCGGA 967  
DB 932 TCAGGAGGCGCACTACTTTCAGATGAGTACCTATTCAACAAACAGAACTATGATACGGA 991  
QY 968 TTCTCTTGGCCTTTGTCTTCCGTGGGAAGAAACAGACACCGGATATTACACCTGCTCTT 1027  
DB 992 TTCTGTTTGTCTTTGTATCATCAGTGGCAAGAAACAGACCGGATACTACACTTTGTTCT 1051  
QY 1028 CTTCAAGACACCCAGCCAGTCAAGCTTTGGTGACCATCTCTAGAAAAGGGTTTATAACG 1087  
DB 1052 CTTCAAGACATCCAGTCAATCAGCTTTGGTTACCATCGTAGGAAGGGATTTATAATG 1111  
QY 1088 CTACCACTCGCAGAGAGATGAAATGACCCGTAGGAAAGTCTGCTTCTCAGTCA 1147  
DB 1112 CTACCAATTCAGTGAAGATGTAATGAAATGACCAATATGAAGATTTGTTTTCTGTCA 1171  
QY 1148 GGTTTAAAGGTCACCCAGCAATCCGATCGACGTGGATCTTCTCAGACCTCATTTCCCT 1207  
DB 1172 GGTTTAAAGCTACCCACAAATCAGATGACGTGGACCTTCTCTCGAAATCATTTCTCT 1231  
QY 1208 GTGAACAGAGAGCTCGAGATGGGTACAGCATATCTAAATTTGGCATCATGAACA 1267  
DB 1232 GTGAGCAAAAGGCTTTGTAAACGGATACAGCATATCCAAGTTTGGCAATCATAGCAC 1291  
QY 1268 AGCCAGGAGATACATATTCTATGCAGAAATGATAGCCCGTACGAAAGTCTTCAACCAATGTCA 1327  
DB 1292 AGCCAGGAGATATATATTCATCGAATAATGATGATGCCCAATTTACCAAAATGTTCA 1351  
QY 1328 CGCTGAATATAAGAAAGAAACCTCAAGTGTGAGCAAAATGCCTCAGCCAGCCAGCGTCT 1387  
DB 1352 CGCTGAATATAAGAAAGAAACCTCAAGTGTGCGCAGAGCATCGCAAGTCAGGCGTCT 1411  
QY 1388 GTTCCTGTATGGCTACCCGTACCCCTTGGACCTGGAAGAAGTGTTCGAGCAAAATCT 1447  
DB 1412 GTTCTCGGATGATACCCATTACCATCTTGGACCTGGAAGAAGTGTTCAGACAAGTCTC 1471  
QY 1448 CCAATTCGAGGAGGAATCCCAAGAGGTTTGGAAATAAAGGCTTAACAGAAAGTGT 1507  
DB 1472 CCAACTGCACAGAAGATCATCAAGAGGCTGGAATAGAAAGGCTTAACAGAAAGTGT 1531  
QY 1508 TTGCCAGTGGGTCTCAGAGAGTACTCTAAATATGAGTGGGCGGGGAAAGGCTTCTGG 1567  
DB 1532 TTGCACAGTGGGTCTCAGAGAGTACTCTAAACATGAGTGAAGCCATAAAGGGTCTCTGG 1591  
QY 1568 TCAATGCTGTGCTACAAATTCATGGGCGAGCTTTCGCAAAACCATCTTTTAACTCAC 1627  
DB 1592 TCAAGTGTGTGCATACAAATTCCTCTTGGCACATCTTGTGAGACGATCCTTTTAAACTCTC 1651

QY 1628 CAGGCCCTTCCCTTTTCATCAAGAACACATCTCTCTTATTCGACCATTTGGCTCTGTC 1687  
DB 1652 CAGGCCCTTCCCTTTTCATCAAGAACACATCTCTCTTATTCGACCAATTTGGTGTGTC 1711  
QY 1688 TCCCTTTCATTTGTTGTTCTCATTTGTTGATGTCGCACAAATACAAAAGCAATTTAGGT 1747  
DB 1712 TCCCTTTCATTTGTTGTTTAAACCTTAAATTTGTACAAAGTACAAAAGCAATTTAGGT 1771  
QY 1748 ACGAGAGTCACTCAGATGATCCAGGTGACGTGCGCCCTGGATACGAGTACTTCTACG 1807  
DB 1772 ATGAAAGCCAGTACAGATGGTACAGGTGACCGGCTCCCTCAGATATGAGTACTTCTACG 1831  
QY 1808 TTGACTTCAGGAGTATGAATATGACCTTAAGTGGAGTTCGCCGAGAGAGAACTTAGAGT 1867  
DB 1832 TTGATTTAGAGAAATATGAATATGATCTCAATGGAGTTTCCAGAGAAATTTAGAGT 1891  
QY 1868 TTGGAAGGTCTCGGGTCTGGCGTTCGGGAGGTGATGAACGCCACCGCTATGGCA 1927  
DB 1892 TTGGAAGGTACTAGGATCAGGTGCTTTTGGAAAAGTATGAACGCAACAGCTTATGGAA 1951  
QY 1928 TTAGTAAACGGGAGTCTCAATTCAGGTGCGGCTGAAGTCTAAAGAGAGAAAGCTTGACA 1987  
DB 1952 TTAGCAAAACAGGAGTCTCAATCCAGGTGCGGCTCAAAATGCTGAAAGAAAGAGAGACA 2011  
QY 1988 GCTGTGAAAAGAAAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047  
DB 2012 GCTCTGAAGAGAGGACACTCATGTCAGAACTCAAGATGATGACCCAGCTGGGAAGCCAG 2071  
QY 2048 ACAACATCTGAATCTGCTGGGGGATGCACACTGTCAGGGCAGTGTACTTGAATTTTG 2107  
DB 2072 AGAATATTGTGAACCTGCTGGGGGCTGCACACTGTCAGGACCAATTTACTTGATTTTG 2131  
QY 2108 AATATTGTTGCTATGCTGACCTCCCTCACTACTCTAAGAGTAAAGAGAGAGATTTTACA 2167  
DB 2132 AATACTGTTGCTATGCTGATCTTCTCACTATCTAAGAGTAAAGAGAGAAATTTTACA 2191  
QY 2168 GGACATGACAGAGATTTTAAAGAAACATAATTTCACTTCTTACCTACTTTCCAGGCAC 2227  
DB 2192 GGACTTGGACAGAGATTTTCAAGGAACACAAATTTCACTTTTACCCACTTTCCATCAC 2251  
QY 2228 ATTCAANTTCCAGCATGCTGTTTCAAGAGAGTTCAGTTACACCGCCCTTGGATCAGC 2287  
DB 2252 ATCCAAATTCAGCATGCTGTTTCAAGAGAGAGTTCAGATACACCGGACTCGGATCAA 2311  
QY 2288 TCTCAGGTTTCAATCGGAATTCAAATTCATCTGAAGATCAGATGATGAATAACAGAGA 2347  
DB 2312 TCTCAGGCTTCATCGGAATTCATTTCACTCTGAAGATGAAATTTGAATATGAARCCAAA 2371  
QY 2348 AGAGCTGGCAGAGAAGAGGAGGAGGATTTGAACGTGCTGACGTTTGAAGACCTCTCTT 2407  
DB 2372 AAAGCTG-----GAGAGAGAGGAGGACTTGAATGCTTACATTTGAAGATCTTCTT 2425  
QY 2408 GCTTTGCTTACCAAGTGGCCAAAGCATGGAATTCCTGGAGTTCAAGTCTGTGTCTTCCACA 2467  
DB 2426 GCTTTGCTATCAAGTTCGCAAGGAATGGAATTTCTGGAATTTAAGTCTGTGTCTTACA 2485  
QY 2468 GAGACCTGGCAGCAGAGATGTTGGTCAACCCAGGGAAGTGGTGAAGATCTGTGACT 2527  
DB 2486 GAGACCTGGCCGAGGAAAGTGTGTCACCCAGGGAAGTGGTGAAGATGATGACT 2545  
QY 2528 TTGGACTGGCCGAGACATCTCTGAGGACTCCAGCTTACCTGTCAGGGGCAACGCGACGG 2587  
DB 2546 TTGGATTTGGCTCGAGATTCATGAGTATTCACATATGTTGTACAGGGCAATGCCGCTC 2605  
QY 2588 TGCCCGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACAAATCAAGAGTG 2647  
DB 2606 TGCTGTGTAATGGATGGCCCGGAGAAAGCTTGTGAAGGCATCTACACATTAAGAGTG 2665  
QY 2648 AGCTCTGCTTACGGCATCTTCTCTCGTGGGAGATATTTTCACTGGGTGTGAACCTTACC 2707  
DB 2666 ATGCTGTGTCATATGGAATATTACTGTGGAAATCTTCTCACTTGGTGTGAATCTTACC 2725  
QY 2708 CTGGCAATCTCTCGACGCTACTTCTATAAACTGATTCAGAGTGGATTTAAATGGAG 2767

Db 2726 CTGCGATCCGGTGTGCTAACTTCTACAACTGATTCRAAATGGATTTAAATGGATC 2785  
 Qy 2768 AGCCATTCTATGCACAGAGGATATACCTTTGTAATGCAATCTCTGCTGGCTTTTGACT 2827  
 Db 2786 AGCCATTCTATGCACAGAGGATATACCTTTGTAATGCAATCTCTGCTGGCTTTTGACT 2845  
 Qy 2828 CAAGGAAGCGCCATCTTCCCAACCTGACCTTCTTTAGGATGTCAGCTGCAGAG 2887  
 Db 2846 CAAGGAAGCGCCATCTTCCCAACCTGACCTTCTTTAGGATGTCAGCTGCAGATG 2905  
 Qy 2888 CAGAAGAAG-----CATGTATCAGAACATCCATCCATCTACC 2924  
 Db 2906 CAGAAGAAGCGATGATCAGAATGCGATGCGCTGTTTCGGAATGCTCTCACACCTACC 2965  
 Qy 2925 AAAACAGCGCCCTCAGCAGAGAGGGGCTCAGAGC-CCAGTGGCCACAGCGCCAGG 2983  
 Db 2966 AAAACAGCGCCCTCAGCAGAGAGATGGATTTGGGGCTACTCTCTCCGAGGCTCAGG 3025  
 Qy 2984 T-GAAGATTACAGAGAAAGTTAGCGAGGAGGCTTTGGACCCCGCC----ACCCTAG 3038  
 Db 3026 TCGAAGATTCTGAGGAGCAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAA 3085  
 Qy 3039 CAGGCTGTAGACCGCAGCAAGATTAGCCTCGCCTCT--GAGGAAGCGCCCTACAGCG 3096  
 Db 3086 CAGGCTGTAGATTACCAAAACAAGATTAAATTCATCACTAAAGAAATCTATTATCAAC 3145  
 Qy 3097 CGTTGCTTCGCTGACTTTCTCTAGATGCTGTGCTGCCATTACTC-----CAAAGTGA 3149  
 Db 3146 TGCTGCTTCACGACACTTTCTCTAGAGCGCTGCTGCTTTACTCTTTTCTTCAAAGGGA 3205  
 Qy 3150 CTTCTATAAATCAAACTCTCTCGCACAGCGGGGAGAGCCAAATAGAGACTGTGTGG 3209  
 Db 3206 CTTTGTAAATCAAACTCTCTGTCACAGGCGAGGAGCTGATAATGAACCTTTATTG 3265  
 Qy 3210 TGAGCGCGCTTACCTTGGGGCCTTCCACAGAGCTTGAGGGAAAGCCATGATCTGAAA 3269  
 Db 3266 GAGCATTGATCTGCATCAAGCCCTCTCAGCGCGCTTGAGTGAATGTGTACCTGAAG 3325  
 Qy 3270 TATGATATTTCTGTAATACGTGAACAACCAACCCGCTTTTCTGCTAAGGGAAGC 3329  
 Db 3326 TACAGTATATCTTGTAAATACATAAAACAAA-----AGCATTTTCTAAGGAGAGC 3378  
 Qy 3330 TAAATATGATTTTAAATAATCTATGTTTAAATACTATGTAACCTTTTCTATCTATTAG 3389  
 Db 3379 TAATATGATTTT--AAGTCTATGTTTAAATAATATGTAATTTTTCACCTATTAG 3436  
 Qy 3390 TGATATATTTATGGATGGAATAAATCTTCTACTGTAAAAAATAAATAAATAAATAA 3449  
 Db 3437 TGATATATTTATGGGTGGATTAATAATTTCTACTACAGAAAAAATAAATAAATAA 3496  
 Qy 3450 AAAA 3453  
 Db 3497 AAAA 3500

## RESULT 14

AAAT72117

ID AAAT72117 standard; cDNA; 3501 BP.

XX AC

XX AAAT72117;

 XX DT 19-AUG-1997 (first entry)  
 XX DE Human flk-2 receptor coding sequence.

XX KW

 KW pTK; liver; fetal liver kinase 2; flk2; receptor protein tyrosine kinase;  
 KW haematopoietic hierarchy; thymus; adult; brain; bone marrow; primitive portion;  
 KW proliferation; differentiation; extracellular domain; soluble form; ligand;  
 KW macrocytic anaemia; aplastic anaemia; cancer; radiation; mouse; ss.  
 XX OS Homo sapiens.

XX Key Location/Qualifiers  
 FH CDS 58..3039  
 FT /\*tag= a  
 FT /product= Human\_flk-2\_receptor  
 FT 58..138  
 FT sig\_peptide /\*tag= b  
 FT 139..3036  
 FT mat\_peptide /\*tag= c  
 FT 139..3036  
 XX US5621090-A.  
 PN 15-APR-1997.  
 XX 02-APR-1991; 91US-0679666.  
 XX 26-JUN-1992; 92US-0906397.  
 PR 02-APR-1991; 91US-0679666.  
 PR 28-JUN-1991; 91US-0728913.  
 PR 15-NOV-1991; 91US-0793065.  
 PR 24-DEC-1991; 91US-0813593.  
 XX (UYPR-) UNIV PRINCETON.  
 PA Lemischka IR;  
 PI WPI; 1997-235228/21.  
 DR P-PSDB; AAW19873.  
 XX Protein containing the extracellular domain of human flk-2 - used  
 PT for identification of primitive haematopoietic cell proliferation  
 PT and differentiation stimulatory ligands, e.g. for treating anaemia  
 XX Claim 1; Fig 1B; 55pp; English.  
 PS This sequence encodes the human fetal liver kinase 2 (flk2). flk-2 is  
 CC a receptor protein tyrosine kinase (ptk) and is important in transducing  
 CC putative self-renewal signals from the environment. flk-2 is expressed  
 CC in fetal liver, spleen, fetal thymus, adult brain and adult bone marrow,  
 CC and it is thought that flk-2 is expressed in the entire primitive portion  
 CC of the haematopoietic hierarchy. The invention concerns a recombinant  
 CC nucleic acid, preferably mRNA, which encodes a protein containing only  
 CC the extracellular domain of human flk-2 and lacking the flk-2 intra-  
 CC cellular catalytic domain. The resultant protein represents a soluble  
 CC form of flk-2 which is used to isolate specific ligands for flk-2. These  
 CC ligands can be used to stimulate proliferation and/or differentiation of  
 CC mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for  
 CC treatment of macrocytic or aplastic anaemia or bone marrow damage caused  
 CC by cancer treatment or radiation.  
 SQ Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;

Query Match 64.5%; Score 2227.2; DB 18; Length 3501;  
 Best Local Similarity 79.6%; Pred. No. 0;  
 Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

Qy 8 GCCTACCGCGCTCGGAGGCGCATCGGCGCTTGGCGAGCGACGCGCGCGCTGC 67  
 Db 35 GGGGACCCCGGCTCCGGAGGCGCATCGGCGCTTGGC---GGCGACGGGGACCGTGC 91  
 Qy 68 TGCTGCTTGTGTTTGTGTCAGTAAATGATTTCTGAGACCGTTACAAACCAAGACCTGCCTG 127  
 Db 92 CGCTGCTGCTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCCTG 151  
 Qy 128 TGATCAAGTGTCTTTTAAATCAGTCATGAGAACAAATGGCTCATCAGCGGGGAAGCCATCAT 187  
 Db 152 TGATCAAGTGTCTTTTAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 211  
 Qy 188 CTTACCGAATGTTGGGAGGATCCCGAGAGACCTCCAGTGTACCCCGGAGCGCCAGAGTG 247  
 Db 212 CATATCCCATGTTATCATCCCGAAGACCTCCGGGTGTGCGTTTGAGACCCAGAGCT 271  
 Qy 248 AAGGACGCTATATGAGCGGCCACCGTGGAGTGGCCGAGTCTGTGGTCTCATCACCTGC 307

1352	CGCTGAATATAAGAGAAACCTCAAGTGTCTCGAGGAAGCANTCGGCAAGTCTACGGCGTCTCT	1411
1388	GTTCCTCTGATGCTACCCCGCTACCCCTCTTTGGACCTGGAAAGAGTGTTCGGACAAATCTC	1447
1412	GTTCCTCGGATGGATACCCATTACCATCTTTGGACCTGGAAAGAGTGTTCAGACAAAGTCTC	1471
1448	CAAATTGCACGGAGGAAATCCCAAGAGGAGTGTGGAAATAAAAGGCTTAACAGAAAAAGTGT	1507
1472	CCAACTGCACAGAAGAGATCACAGAAGAGTGTGGAATAGAAAAGGCTTAACAGAAAAAGTGT	1531
1508	TTGGCCAGTGGGTGTGAGCAGTACTCTAAATATGAGTGAAGCGCGGGAAGGCTTCTCGG	1567
1532	TTGGACAGTGGGTGTGAGCAGTACTCTAAATGAGTGAAGCGCTTAAGAGGTTCTCTCGG	1591
1568	TCAAATGCTGTGGTACAATTTATGGGACGCTTTGCGAAACCATCTTTTAAACTCAC	1627
1592	TCAAGTGTGTGCATACAATTTCCCTTGGCACATCTTGTGAGACGATCTTTTAAACTCTC	1651
1628	CAGGCCCTTCCCTTTATCCCAAGACACATCTCTTCTATGGACCATTTGGGCTCTGTC	1687
1652	CAGGCCCTTCCCTTTATCCCAAGACACATCTCTTCTATGCAACAATTTGGTGTGTC	1711
1688	TCCCCTTCAATGTGTCTCAATTTGTGTGATCTGCCACAAATACAAAAGCAATTTAGGT	1747
1712	TCCTCTTCAATGTGTCTCAATTTGTGTGATCTGCCACAAATACAAAAGCAATTTAGGT	1771
1748	ACGAGAGTCACTGTCAGATGATCCAGTGACTGGCCCCCTTGGATTAACGAGTACTTCTACG	1807
1772	ATCAAGCCAGCTACAGATGGTACAGGTGACCGCTCTCAGATAATGACTACTTCTACG	1831
1808	TTGACTTTCAGGGACTATGAATATGACCTTAAAGTGGGAGTTCGCGAGAGAACTTTAGAGT	1867
1832	TTGATTTTCAGAGAAATATGAATATGATCTCAATTTGGGAGTTCCTCAAGAGAAAAATTTAGAGT	1891
1868	TTGGGAAGTCTCTGGGCTGTGGCGCTTTCGGGAGGATGATGAACGCCACGCCCTATGGCA	1927
1892	TTGGGAAGTCTCTGGGCTGTGGCGCTTTCGGGAGGATGATGAACGCCACGCCCTATGGCA	1951
1928	TTAGTAAACGGGAGTCTCAATTCAGGTGGCGTGAAGATGCTTAAAGAGAAAGCTGACA	1987
1952	TTAGCAAAACAGAGTCTCAATCCAGGTTCGCGTCAAAATGCTTCAAGAAAAAGCAGACA	2011
1988	GCTGTGAAAAGAGTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGACACCAGT	2047
2012	GCCTCTAAAGAGAGGACCTCATGTCAAGACTCAAGATGATGACCCAGCTGGGAAGCCAGC	2071
2048	ACAACATCTGTAATCTGCTGGGGGATGCACACTGTGAGGCCAGTGTACTTGTATTTTG	2107
2072	AGAATATTGTGAACCTGTGGGGGCTGCACACTGTGAGGCCAGTGTACTTGTATTTTG	2131
2108	AATATTGTGTATGTGTGACTCTCAACTTACCTAAGAGTAAAGAGAGAAAGTTTTCACA	2167
2132	AATACTGTGTATGTGTGATCTTCTCACTATCTAAGAGTAAAGAGAAAAATTTTACA	2191
2168	GGACATGGACAGAGATTTTAAAGAACATAAATTTTCACTTCTTACCCTTACTTTTCCAGGCAC	2227
2192	GGACTTTGGACAGAGATTTTCAAGGAACACAATTTTCACTTCTTACCCTTACTTTTCCAAATCAC	2251
2228	ATTCAAATTTCCAGCATCTGGTGTTCAGGAGAGTTCAGTTTACACCCGCCCTTGGATCAGC	2287
2252	ATCCAAATTTCCAGCATCTGGTGTTCAGGAGAGTTCAGTTTACACCCGCCCTTGGATCAGC	2311
2288	TCTCAGGTTTCAATGGGAATTTCAATTTCTCAAGATGAGTTGAATTTGAATAAACCCAGA	2347
2312	TCTCAGGCTTCATGGGAATTTCAATTTTCACTCTCAGATGAAATTTGAATTTGAATAAACCCAAA	2371
2348	AGAGGCTGCGCAGAAAGAGGAGGAGAAATTTTGAACCTGTGACGTTTGAAGACCTCTTT	2407
2372	AAAGGCTG-----GAAGAGAGAGGAGCTTGAATGTGCTTACATTTGAAGATCTTCTTT	2425
2408	GCCTTTCGCTACCAAGTGGCCAAAGGAGGAGGAGTTCCTGGAGTTCAAGTCTGTGTCTCACA	2467
2426	GCCTTTCGATATCAAGTTGCAAGAGGAGGAGTTCCTGGAGTTTGAATTTGAAGTCTGTGTCTACA	2485

```
QY 2468 GAGACCTGCGCAGCCAGGAATGTGTGTCACCCACCGGAAGTGTGAAGATCTGTGACT 2527
DB 2486 GAGACCTGCGCCGCGCAGACGTGCTGTCTACCCACGGAAGTGTGAAGATATGTGACT 2545
QY 2528 TTGGAGTGGCCGCGAGACATCTTGAGCGACTCCAGCTACCTGTCAGGGCAACGCGACGCG 2587
DB 2546 TTGGAGTGGCTCGAGATATCATGAGTGATTCGAAGTATTCGAAGTATGTGTGTCAGGGCAATGCCGCTC 2605
QY 2588 TGCCGCTGAAGTGGGACCGGAGAGCTTATTTGAAGGGATCTACAAATCAAGATG 2647
DB 2606 TGCCGCTGAAGTGGGACCGGAGAGCTTATTTGAAGGGATCTACAAATCAAGATG 2665
QY 2648 AGCTGTGCTCCTACGGCATCTCTCTGGGAGATATTTTCAGTGGTGTGAACCCCTTACC 2707
DB 2666 ATGTCTGGTCATATGAATATTAATCTGTGGGAATCTTCTCACTTGTGTGAATCCTTACC 2725
QY 2708 CTGGCAATCTCTGCGAGCTTAATCTATATAACTGATTCAGAGTGGATTTAAATGAGC 2767
DB 2726 CTGGCAATCTCTGCGAGCTTAATCTATATAACTGATTCAGAGTGGATTTAAATGAGC 2785
QY 2768 AGCCATCTTATCCACAGAGGATATATTTGTATGCAATCTCTGCGGCTTTTGACT 2827
DB 2786 AGCCATCTTATCTACAGAGAAATATACATTAATGCAATCTCTGCGGCTTTTGACT 2845
QY 2828 CAAGGAAGCGGCATCTTCCCAACCTGACTTCTATTTTAGTATGCTGAGCTGGCAGAGG 2887
DB 2846 CAAGGAAGCGGCATCTTCCCAACCTGACTTCTATTTTAGTATGCTGAGCTGGCAGAGG 2905
QY 2888 CAGAAGAAG-----CATGTATCAGAAATCCATCATCTACC 2924
DB 2906 CAGAAGAAGCGATGTATCAGAAATGTCGATGGCGCTGTTTCGGAATCTTCTCAACACCTACC 2965
QY 2925 AAACAGAGCGGCGCTCAGCAGAGCGGCTCAGAGC-CCAGTCCGCCACAGCGCCAGG 2983
DB 2966 AAACAGAGCGGCGCTTTCAGCAGAGAGATGGATTTGGGCTACTCTCTCCGCGAGGCTCAGG 3025
QY 2984 T-GAAGATTTCACAGAAAGAAAGTTAGCAGGAGCGCTTGGAGCCCGCC-----ACCCCTAG 3038
DB 3026 TCGAAGATTTCGAGAGAAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAA 3085
QY 3039 CAGGCTGTAGCGCAGAGCCAGATTTAGCTGCGCTCT--GAGGAAGCGCCCTCAGCG 3096
DB 3086 CAGGCTGTAGATTACCAAAACAGATTAATTTTCATCACTAAAGAAATCTATTATCAAC 3145
QY 3097 CGTGTCTGCTGCGACTTTCTCTAGATGCTGCTGCCATTACTC-----CAAGTGA 3149
DB 3146 TGTCTGCTTCACAGACTTTTCTCTAGAACCGCTCTGCGTTTACTCTTTGTTTCAAAGGGA 3205
QY 3150 CTCTATAAAATCAAAACCTCTCTCTCAGAGCGGGGAGAGCCCAATATGAGACTTGTGG 3209
DB 3206 CTCTATAAAATCAAAACCTCTCTCTCAGAGCGGGGAGAGCTGATAATCACTTTATTG 3265
QY 3210 TGAGCCCGGCTACCTCGGGGCGCTTTCAGAGAGCTTGAGGGGAAAGCCATGTATCTGAAA 3269
DB 3266 GAGCAATGTATCTGCATCCAGGCGCTCTCAGGCGGCTTGAGTGAATTTGTACCTCAAG 3325
QY 3270 TATAGTATATCTGTGTAATACGTGAACAAACCAACCCGTTTGTGCTAAAGGAAAGC 3329
DB 3326 TACAGTATATCTGTGTAATACGTGAACAAACCAACCCGTTTGTGCTAAAGGAAAGC 3378
QY 3330 TAAATATGATTTTAAATACTATGTTTAAATACTATGTAATCTTATCTATTTAG 3389
DB 3379 TAAATGATTTT--AGTCTATGTTTAAATACTATGTAATCTTATCTATTTAG 3436
QY 3390 TGATATATTTATGATGGAATAATACTTTCTACTGTAAAAAATAAAAAAAAAAAAAA 3449
DB 3437 TGATATATTTATGCGTGAATAATACTTTCTACTACAGAAAAAATAAAAAAAAAAAAAA 3496
QY 3450 AAAA 3453
DB 3497 AAAA 3500
```

```
RESULT 15
AAX77515
ID AAX77515 standard; cDNA; 3501 BP.
XX
AC AAX77515;
XX
DT 05-AUG-1999 (first entry)
XX
DE Human flk-2 cDNA.
XX
KW Human; flk-2; cell isolation; fetal liver kinase; receptor;
KW monoclonal; polyclonal; antibody; tyrosine kinase; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 58..3039
FT /tag= a
FT /product= "flk-2"
XX
PN US5912133-A.
XX
PD 15-JUN-1999.
XX
PF 10-FEB-1998; 98US-0021324.
XX
PR 19-NOV-1992; 92US-0977451.
PR 02-APR-1991; 91US-0679666.
PR 28-JUN-1991; 91US-0728913.
PR 15-NOV-1991; 91US-0793065.
PR 24-DEC-1991; 91US-0813593.
PR 26-JUN-1992; 92US-0906397.
PR 12-NOV-1992; 92US-0975049.
PR 30-APR-1993; 93US-0055269.
PR 31-OCT-1994; 94US-0252498.
PR 15-FEB-1996; 96US-0601891.
XX
PA (UYPK-) UNIV PRINCETON.
XX
PI Lemischka IR;
XX
DR WPI; 1999-357194/30.
DR P-PSDB; AAY08617.
XX
PT Isolating hematopoietic cells expressing fetal liver kinase 1
PT receptors
XX
PS Disclosure; Fig 1b; 59pp; English.
XX
CC This invention describes a novel method of isolating cells expressing
CC fetal liver kinase 1 (flk-1) receptors on their surface and comprises
CC binding the cells to a polyclonal or monoclonal antibody specific to
CC the Flk-1 receptor and isolating the cells that have bound to the
CC antibody. The method can be used to isolate hematopoietic stem cells in
CC any mammal but preferably a rat, mouse, rabbit or human. The proteins of
CC the invention belong to the receptor protein family. This sequence
CC encodes the human flk-2 protein which is used in the method of the
CC invention.
XX
SQ Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;
Query Match 64.5%; Score 2227.2; DB 20; Length 3501;
Best Local Similarity 79.68; Pred. No. 0;
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;
QY 8 GGCTACCGCGCTCCCGAGGCCATCGGCGCTTGGCGCAGCGACGCGCGGCTGC 67
DB 35 GGGGACCCCGGCTCCGAGGCCATCGGCGCTTGGCGCAGCGACGCGCGGCTGC 91
QY 68 TGCTGCTGTTGTTTGTTCAGTATGATTTCTGAGACCGGTTTACAAACAGACCTGCTG 127
DB 92 CGTGTCTGCTGTTTGTTCAGTATGATTTCTGAGACCGGTTTACAAACAGATCTGCTG 151
```

QY 128 TGATCAAGTGTGTTTTTAATCAGTCTATGAGAACAAATGGCTCATCAGCGGGAAGGCCATCAT 187  
DB 152 TGATCAAGTGTGTTTTTAATCAGTCTATGAGAACAAATGATTCATCATCTGGGAGTCTATCAT 211  
QY 188 CGTACCGAATGGTGGGAGGATCCCCAGAGACCTCCAGTGTACCCTGGGAGGCCCGCAGAGTG 247  
DB 212 CATATCCCATGGTATCAGAAATCCCGGAGAGCTCGGGTGTGCGTTGAGACCCCGAGAGCT 271  
QY 248 AAGGACCGGTATATGAAGCGCCCGTGGAGGTGGCCGAGTCTGGGTCCATCACCCTGCG 307  
DB 272 CAGGACAGTGTACGAAGCTGCGCGTGTGGAGTGGATGTATCTGCTTCCATCACACTGC 331  
QY 308 AAGTGCAGCTGCCACCCAGCGGACCTTTCCTGCTCTGGGTCTTTAAGCACAGCTCCC 367  
DB 332 AAGTCTGGTGCATGCCCGGAGACATTTCTCTGCTCTGGGTCTTTAAGCACAGCTCCC 391  
QY 368 TGGGCTGCCAGCGGCACCTTTGATTTACAAACAGAGGAATGTTTCCATGGCCATCTTGA 427  
DB 392 TGAATTGCCAGCCACATTTGATTTACAAACAGAGGAGTTGTTCCATGGTCAATTTGA 451  
QY 428 ACGTACAGAGACCCAGCAGAGAAATACCTACTCCATATTCAGAGCGAAGCCGCAACT 487  
DB 452 AAATGACAGAAACCCAGCTGGAGAAATACCTACTTTTATTTCAGAGTGAAGCTACCAATT 511  
QY 488 ACACAGTACTGTTCAGTGAATGTAAAGAGATACACAGCTGTATGTCTAAGGAGACCTT 547  
DB 512 ACACAAATATTGTTACAGTGTAGTATAGAAATACCTGCTTTACACATTAAGAAGACCTT 571  
QY 548 ACTTTAGGAAGTGAAGAACCCAGGATGCACCTCTCTGCATCTCCGAGGGTGTTCGGAGC 607  
DB 572 ACTTTAGAAAATGAAAACCCAGGACGCCCTGGTCTGCATATCTCAGAGCGTTCCAGAGC 631  
QY 608 CCACTGTGGAGTGGTGTCTGCAGCTCCACAGGAGAAAGCTGTAAAGAAAGAGGCCCTG 667  
DB 632 CGATCGTGAATGGTGTGTTGCGATTCACAGGGGGAAGCTGTAAAGAAAGAAAGTCCAG 691  
QY 668 CTGTTGTGAGAAAGGAGGAAAAGTGTTCATGAGTGTGTCGGAACAGACATCAGATGCT 727  
DB 692 CTGTTGTTAAAGAGGAGAAAAGTGTTCATGAATATTGTTGGAGCGGACATGAAGTGTCT 751  
QY 728 GTGCTAGAAATGCACTGGGCGCGAATGCACCAAGCTGTTCACCATAGATCTAAACCCAGG 787  
DB 752 GTGCCAGAAATGAATGGGCGGGAATGCACCCAGGCTGTTCACATAGATCTAAATCAAA 811  
QY 788 CTCCTCAGAGCACATGCCCCAGTATTTCCTGAAAGTGGGGAACCCCTTGTGGATCAGGT 847  
DB 812 CTCCTCAGACCAATTTGCCACAATTTCTTTAAAGTAGGGGAACCCCTTATGGATAAGGT 871  
QY 848 GTAAGGCCATCCATCTGAACCATGATTCGGGCTCACCTGGGAGCTTGGGAAGCAAGCCC 907  
DB 872 GCAAAAGCTGTTCTATGTGAACCATGGATTCGGGCTCACCTGGGAATTAGAAAACAAGCAC 931  
QY 908 TGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTTCCACAAACAGGACCATGATTCGGA 967  
DB 932 TCGAGGAGGCGAACCTACTTTGAGATGAGTACCTATTTCACAAACAGAACTATGATACGGA 991  
QY 968 TTCTCTTGGCCCTTTGTCTTCCTGGGGAAGGAAGACACCCGGATATTAACCTGCTCTT 1027  
DB 992 TTCTCTTGTCTTTGATCATCAGTGGCAAGAAAGCACACCCGGATACTACCTTGTCTCT 1051  
QY 1028 CCTCAAGCACCCAGCCAGTACGCTGCTGACCATCTAGAAAAGGTTTATAAAGC 1087  
DB 1052 CTTCAAAAGCATCCAGTCAATCAGCTTTGGTTTACCATCTGAGAAAGGATTTATAAATG 1111  
QY 1088 CTACAGCTCGCAAGAAGAGTATGAATTTGACCCGCTACGAAAGTCTCTCTCTCAGTCA 1147  
DB 1112 CTACCAATTCAGTGNAGATTTAGAAATTTGACCAATATGAAGATTTTGTCTCTGTC 1171  
QY 1148 GGTTTAAAGCGTACCACCAAGATCCGATGCAGCTGGATCTTCCTCAAGCCCTCATTTCCCTT 1207  
DB 1172 GGTTTAAAGCCCTACCACCAAAATCAGATGTACGTGGACCTTCTCTCGAAAATCATTTCCCTT 1231

QY 1208 GTGAACAGAGAGCGCTGGAGGATGGGTACAGCATATCTAAATTTTCCGATCATATAAGACA 1267  
DB 1232 GTGACGCAAAAGGCTCTTGATAACGGATACAGCATATCCAAGTTTTTGCATCATAGAACC 1291  
QY 1268 ACCGAGGAGATACATATCTATGAGAAAANTGATGACGCCCGCAGTTCCACCAATGTTCA 1327  
DB 1292 AGCCAGGAGAAATATATTTCCATGAGAAAANTGATGATGCCAATTTTACCAAAATGTTCA 1351  
QY 1328 CCTCAATATAAAGAAACCTCAAGTCTAGCAAAATGCTTCAGCCAGCCAGCCGCTCCCT 1387  
DB 1352 CCTCAATATAAGAGAAACCTCAAGTCTGCGAAGCATGCGGAAGTCAAGCGCTCCT 1411  
QY 1388 GTTCTCTGATGGCTACCCGCTACCCCTCTTGGACCTGGGAAGAGTGTTCGGACAATCTC 1447  
DB 1412 GTTCTCGGATGATACCCATTTACCATCTTGGACCTGGGAAGAGTGTTCAGACAAGTCTC 1471  
QY 1448 CCAATTCGACGAGGAAATCCCGAGAGGATTTGGATTAAGAGGCTTAACAGAAAGTGT 1507  
DB 1472 CCNACTGTCACAGAAGAGATCACAGAAGGAGTCTGGAAATAGAAAGGCTTACAGAAAAGTGT 1531  
QY 1508 TTGGCAGTGGGTGTGCGAGCAGTACTTAATATGAGTGAAGCGCGGAAAAGGCTTCTG 1567  
DB 1532 TTGGACAGTGGGTGTGCGAGCAGTACTTAACATGAGTGAAGCCATAAAGGTTCTCTG 1591  
QY 1568 TCAATGTCTGTGCTACAATTTCTATGGGCACGCTTTTGGAAAACCATCTTTTAAACTCAC 1627  
DB 1592 TCAAGTGTCTGCATACAATTTCCCTTGGCACATCTTTGTGAGACGATCTTTTAACTCTC 1651  
QY 1628 CAGGCCCCCTTCCCTTATCCAAAGACAATCTCTTATGCGACCATTTGGGCTCTGTC 1687  
DB 1652 CAGGCCCCCTTCCCTTATCCAAAGACAATCTCTTATGCAACAATTTGGTGTGTC 1711  
QY 1688 TCCCTTCACTTGTGTTCTCATTTGTGATGTCGCACAAATACAAAAGCAATTTAGGT 1747  
DB 1712 TCTCTTCAATTCGTTTTTAAACCTGCTAATTTGTCAAGTACAAAAGCAATTTAGGT 1771  
QY 1748 ACAGAGTCAAGTGCAGATGATCCAGGTGACTGGCCCCCTGGATACAGAGTACTTCTACG 1807  
DB 1772 ATGAAGCCAGCTACAGATGGTACAGGTGACCGCTCTCAGATATAGTACTTCTACG 1831  
QY 1808 TTGACTTTCAGGACTATGAATATGACCTTAAAGTGGGAGTTTCCGAGAGAACTTAGGT 1867  
DB 1832 TTGATTTACAGAAATATGAATATGATCTCAAAATGGGAGTTTCCAAGAGAAAATTTAGGT 1891  
QY 1868 TTGGGAGTCTCTGGGCTCTGGGCTTTCGGGAGGATGATGAACGCCACCCCTATGSCA 1927  
DB 1892 TTGGGAGTCTTAGGATCAGGTGCTTTGGAAAAGTGTGAACGCAACAGCTTATGGA 1951  
QY 1928 TTAGTAAAACGGGAGTCTCAATTTCAAGTGGCGGTGAAGATGCTAAAAGAGAAAGCTGACA 1987  
DB 1952 TTAGCAAAACAGAGTCTCAATCCAGGTGCGCTCAAAATGCTGAAAAGAAAAGCAGACA 2011  
QY 1988 GTGTGAAAAGAGAGTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047  
DB 2012 GCTCTGAAAAGAGGAGCTCATGTCAGAACTCAAGATGATGACCCAGCTGGGAGGCCAG 2071  
QY 2048 ACAACATCTGATCTCTGGGCGCATGCACCTGTCAAGGCGGAGTGTCTGATTTTGTG 2107  
DB 2072 AGAATATTGTGAACCTGCTGGGGGCGTGCACATGTCAGGACCAATTTACTGATTTTGTG 2131  
QY 2108 ATATTTGTTGCTATGCTGACCTCCTCAACTTACCTAAGAAGTAAAAGAGAGAAAGTTTCA 2167  
DB 2132 AATACTGTTGCTATGCTGATCTTCTCACTATCTAAGAAGTAAAAGAGAAAATTTTCA 2191  
QY 2168 GGACATGGACAGAGATTTTAAAGAACATAAATTTCAAGTCTTACCTTACTTTTCCAGCAC 2227  
DB 2192 GGACTTGGACAGAGATTTTCAAGGAACACAATTTCAAGTCTTACCTTACTTTTCAATCAC 2251  
QY 2228 ATTCAAATTCAGCATGCTGTTTACAGGAAAGTTCAGTTTACACCCCGCTTGGATCAGC 2287  
DB 2252 ATCCAAATTCAGCATGCTGTTTCAAGGAAGTTCAGATACACCCCGGACTTCGATCAAA 2311  
QY 2288 TCTCAGGGTTCAATGGGAATTCATTTCTCAAGATGAGATTTGAATATGAAAACCCAGA 2347





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 04:23:58 ; Search time 144.496 Seconds  
(without alignments)  
7328.626 Million cell updates/sec

Title: US-09-919-408-1  
Perfect score: 3453  
Sequence: 1 GCGGCTGGCTACCGCGGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3453	100.0	3453	1	US-07-813-593-1
2	3453	100.0	3453	1	US-07-977-451-1
3	3453	100.0	3453	1	US-07-946-507-1
4	3453	100.0	3453	1	US-08-252-517-1
5	3453	100.0	3453	1	US-07-906-397A-1
6	3453	100.0	3453	1	US-08-601-891-1
7	3453	100.0	3453	5	US-09-021-324-1
8	3453	100.0	3453	5	US-09-021-324-1
9	3453	100.0	3453	5	US-09-021-324-1
10	3453	100.0	3453	5	US-09-021-324-1
11	3453	100.0	3453	5	US-09-021-324-1
12	3453	100.0	3453	5	US-09-021-324-1
13	3453	100.0	3453	5	US-09-021-324-1
14	3453	100.0	3453	5	US-09-021-324-1
15	3453	100.0	3453	5	US-09-021-324-1
16	3453	100.0	3453	5	US-09-021-324-1
17	3453	100.0	3453	5	US-09-021-324-1
18	3453	100.0	3453	5	US-09-021-324-1
19	3453	100.0	3453	5	US-09-021-324-1
20	3453	100.0	3453	5	US-09-021-324-1
21	3453	100.0	3453	5	US-09-021-324-1
22	3453	100.0	3453	5	US-09-021-324-1
23	3453	100.0	3453	5	US-09-021-324-1
24	3453	100.0	3453	5	US-09-021-324-1
25	3453	100.0	3453	5	US-09-021-324-1
26	3453	100.0	3453	5	US-09-021-324-1
27	3453	100.0	3453	5	US-09-021-324-1

28	2081.8	60.3	3120	5	PCT-US95-04228-22	Sequence 22, Appl
29	1206	34.9	1894	4	US-07-912-122-3	Sequence 3, Appl
30	1206	34.9	1894	5	PCT-US93-06404-3	Sequence 3, Appl
31	329	9.5	5084	1	US-08-306-691B-21	Sequence 21, Appl
32	329	9.5	5084	5	PCT-US93-06251-25	Sequence 25, Appl
33	286.4	8.3	4054	1	US-08-180-195-35	Sequence 35, Appl
34	286.4	8.3	4054	1	US-08-477-329-35	Sequence 35, Appl
35	286.4	8.3	4054	2	US-08-475-458-35	Sequence 35, Appl
36	286.4	8.3	4054	3	US-08-980-400-35	Sequence 35, Appl
37	286.4	8.3	4054	4	US-09-583-459A-35	Sequence 35, Appl
38	286.4	8.3	4054	4	US-09-583-210-35	Sequence 35, Appl
39	286.4	8.3	4054	4	US-09-583-449A-35	Sequence 35, Appl
40	286.4	8.3	4054	4	US-09-435-059-35	Sequence 35, Appl
41	286.4	8.3	4100	1	US-08-168-917-3	Sequence 3, Appl
42	286.4	8.3	4100	2	US-08-460-510-3	Sequence 3, Appl
43	286.4	8.3	4100	2	US-08-460-490-3	Sequence 3, Appl
44	286.4	8.3	4100	5	PCT-US92-00730-3	Sequence 3, Appl
45	286.4	8.3	4100	5	PCT-US92-00862-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1  
US-07-813-593-1  
Sequence 1, Application US/07813593  
Patent No. 5185438  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMCLONE SYSTEMS INCORPORATED  
STREET: 180 VARICK STREET  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION NUMBER: US/07/813,593  
FILING DATE: 19920415  
CLASSIFICATION: 435  
PRIOR APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-PPP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3453 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS

```

; LOCATION: 31...3009
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 31...3006
US-07-813-593-1

Query Match      100.0%; Score 3453; DB 1; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGCTGGCTACCGCGGCTCCGAGGCGCATCGGGGGTGGCGCAGCGGCGAGCGG 60
Db 1 GCGGCTGGCTACCGCGGCTCCGAGGCGCATCGGGGGTGGCGCAGCGGCGAGCGG 60

Qy 61 GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db 61 GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120

Qy 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCAATGAGACAAATGGCTCATCAGCGGAAAG 180
Db 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCAATGAGACAAATGGCTCATCAGCGGAAAG 180

Qy 181 CCATCATGTCACCGAATGGTGGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGCGC 240
Db 181 CCATCATGTCACCGAATGGTGGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGCGC 240

Qy 241 CAGAGTGAAGGACGGTATATCAAGCGGCCACCGTGGAGGTGGCGAGTCTGGGTCCATC 300
Db 241 CAGAGTGAAGGACGGTATATCAAGCGGCCACCGTGGAGGTGGCGAGTCTGGGTCCATC 300

Qy 301 ACCCTGCAAGTGCAGTCCCGAGCGGACCTTTGATTTACAAAACAGAGGAATCGTTCCATGGCC 420
Db 301 ACCCTGCAAGTGCAGTCCCGAGCGGACCTTTGATTTACAAAACAGAGGAATCGTTCCATGGCC 420

Qy 421 ATCTTGAAGCTGACAGAGACCCAGGAGGAGAAATACCTACTCATATTCAGAGCGAAGCG 480
Db 421 ATCTTGAAGCTGACAGAGACCCAGGAGGAGAAATACCTACTCATATTCAGAGCGAAGCG 480

Qy 481 GCCAACTACACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGCTAAGG 540
Db 481 GCCAACTACACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGCTAAGG 540

Qy 541 AGACCTTACTTTAGGAGATGAAACACAGGAGGAGAAATACCTACTCATATTCAGAGCGAAGCG 600
Db 541 AGACCTTACTTTAGGAGATGAAACACAGGAGGAGAAATACCTACTCATATTCAGAGCGAAGCG 600

Qy 601 CCGGAGCCCACTGTGAGTGGTGTCTGCAGCTCCAGGAGGAAAGCTGTAAGAGAA 660
Db 601 CCGGAGCCCACTGTGAGTGGTGTCTGCAGCTCCAGGAGGAAAGCTGTAAGAGAA 660

Qy 661 GGGCTGCTGTTGTCAGAAAGGAGGAAAGGTAATCTCATGAGTGTTCGGAACAGACATC 720
Db 661 GGGCTGCTGTTGTCAGAAAGGAGGAAAGGTAATCTCATGAGTGTTCGGAACAGACATC 720

Qy 721 AGATGCTGTCTAGAAATGCACTGGGCGGGAATGCAACCAAGCTGTTCACCATAGATCTA 780
Db 721 AGATGCTGTCTAGAAATGCACTGGGCGGGAATGCAACCAAGCTGTTCACCATAGATCTA 780

Qy 781 AACCAGGCTCCTCAGAGCACACTGCCCCAGTTATCTGAACTGGGGAAACCTTGTGG 840
Db 781 AACCAGGCTCCTCAGAGCACACTGCCCCAGTTATCTGAACTGGGGAAACCTTGTGG 840

Qy 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGGAATTCGGGCTCACCTGGAGCTGGAAGAC 900
Db 841 ATCAGGTGTAAGGCCATCCATGTGAACCATGGAATTCGGGCTCACCTGGAGCTGGAAGAC 900

Qy 901 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTCTCCACAAACAGGACCATG 960
Db 901 AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTCTCCACAAACAGGACCATG 960
```

2041 CACCATGACACATCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCAGGTACTTG 2100  
2041 CACCATGACACATCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCAGGTACTTG 2100  
2101 ATTTTGAATATGTTGCTATGCTGACCTCCCACTACCTAAGAAAGTAAAGAGAGAAG 2160  
2101 ATTTTGAATATGTTGCTATGCTGACCTCCCACTACCTAAGAAAGTAAAGAGAGAAG 2160  
2161 TTTCACAGACATGACACAGAGATTTTAAAGAACATAATTTTACCTACTTTTC 2220  
2161 TTTCACAGACATGACACAGAGATTTTAAAGAACATAATTTTACCTACTTTTC 2220  
2221 CAGGACACATTTAAATTCAGCATGCTGCTTACGAGAAAGTTCAGTTACACCGCCCTTG 2280  
2221 CAGGACACATTTAAATTCAGCATGCTGCTTACGAGAAAGTTCAGTTACACCGCCCTTG 2280  
2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTTCAATCTGGAAGATGAGATTGAATGAA 2340  
2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTTCAATCTGGAAGATGAGATTGAATGAA 2340  
2341 AACCAAGAGGCTGGCAG 2400  
2341 AACCAAGAGGCTGGCAG 2400  
2401 CTCCTTTGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460  
2401 CTCCTTTGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460  
2461 GTCCACAGAGACCTGGCAGCAGGAGATGTTGGTCAACCCAGGGAAGTGTGAAGATC 2520  
2461 GTCCACAGAGACCTGGCAGCAGGAGATGTTGGTCAACCCAGGGAAGTGTGAAGATC 2520  
2521 TGTGACTTTGACTGGCCGCGAGACATCTGACGAGTCTGACGAGTCTGACGAGGCAAC 2580  
2521 TGTGACTTTGACTGGCCGCGAGACATCTGACGAGTCTGACGAGTCTGACGAGGCAAC 2580  
2581 GCACGGCTGCGGCTGAGTGGATGACCCGAGAGCTTATTTGAAGGGATCTACACAATC 2640  
2581 GCACGGCTGCGGCTGAGTGGATGACCCGAGAGCTTATTTGAAGGGATCTACACAATC 2640  
2641 AAGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700  
2641 AAGAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700  
2701 CCTTACCTGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
2701 CCTTACCTGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760  
2761 ATGGAGCAGCCTTCTATGCCACAGAGGATATCTTTGTAATGCAATCTGCTGGGCT 2820  
2761 ATGGAGCAGCCTTCTATGCCACAGAGGATATCTTTGTAATGCAATCTGCTGGGCT 2820  
2821 TTTGACTCAAGGAAGCGGCTTCTTCCCAACCTGACTTCAATTTTAGGATGTCAGCTG 2880  
2821 TTTGACTCAAGGAAGCGGCTTCTTCCCAACCTGACTTCAATTTTAGGATGTCAGCTG 2880  
2881 CGAGAGCAGAGAGAGATGATACAGACATCCATCCATCTACCAAAACAGCGGCCCT 2940  
2881 CGAGAGCAGAGAGAGATGATACAGACATCCATCCATCTACCAAAACAGCGGCCCT 2940  
2941 CAGCAGAGAGGGGCTCAGAGCCAGTGCACAGCGCAGGTGAAGATTCACAGAGAA 3000  
2941 CAGCAGAGAGGGGCTCAGAGCCAGTGCACAGCGCAGGTGAAGATTCACAGAGAA 3000  
3001 AGAAGTTAGCAGAGGGCTTGGAGCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCA 3060  
3001 AGAAGTTAGCAGAGGGCTTGGAGCCCGCCACCTAGCAGGCTGTAGACCGCAGAGCA 3060  
3061 AGATTAGCCTGCTGCTGAGGAGCGCCCTACAGCGGCTGCTGCTGCTGCTGCTGCTGCT 3120  
3061 AGATTAGCCTGCTGCTGAGGAGCGCCCTACAGCGGCTGCTGCTGCTGCTGCTGCTGCT 3120  
3121 AGATGCTGCTGCCATTTACTCCAAAGTGAATCTTCTATAAATCAAACTCTCTCGCACAG 3180

Db 3121 AGATGCTGCTGCCATTTACTCCAAAGTGAATCTTCTATAAATCAAACTCTCTCGCACAG 3180  
QY 3181 GCGGAGAGCAATAATAGACTGTGTTGGTGAGCCGCCCTACCTGGGGGCTTTTCCACG 3240  
Db 3181 GCGGAGAGCAATAATAGACTGTGTTGGTGAGCCGCCCTACCTGGGGGCTTTTCCACG 3240  
QY 3241 AGCTTCAGGGGAGGCAATGATCTGAAATATAGTATATTTCTGTAATACGTGAACAA 3300  
Db 3241 AGCTTCAGGGGAGGCAATGATCTGAAATATAGTATATTTCTGTAATACGTGAACAA 3300  
QY 3301 ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTAAAAATCTATCTTTAA 3360  
Db 3301 ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTAAAAATCTATCTTTAA 3360  
QY 3361 AATACTATGTAACCTTTTTCATCTATTTAGTATATATTTTATGATGGAATAAATCTTC 3420  
Db 3361 AATACTATGTAACCTTTTTCATCTATTTAGTATATATTTTATGATGGAATAAATCTTC 3420  
QY 3421 TACTGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3453  
Db 3421 TACTGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3453

RESULT 2  
US-07-977-451-1  
; Sequence 1, Application US/07977451  
; Patent No. 5270458  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/977,451  
; FILING DATE: 19921119  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US UNASSIGNED  
; FILING DATE: 12-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3453 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 112..3006
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 31..111
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..3009
; US-07-977-451-1

Query Match 100.0%; Score 3453; DB 1; Length 3453;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGTGGCTACCGCGCGCTCCGGAGCCATCGCGGCGTTGGCGAGCGCGACGCCG 60
DB 1 GCGGCGTGGCTACCGCGCGCTCCGGAGCCATCGCGGCGTTGGCGAGCGCGACGCCG 60

QY 61 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 61 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120

QY 121 CTGCGCTGTCATCAAGTGTCTTTAATCAGTCATGAGAACAAATGGCTCATCGCGGAAAG 180
DB 121 CTGCGCTGTCATCAAGTGTCTTTAATCAGTCATGAGAACAAATGGCTCATCGCGGAAAG 180

QY 181 CCATCATCTGATCGGATGCGGAGGATCCCGAGAAAGCTCCAGTGTACCCCGAGGCGC 240
DB 181 CCATCATCTGATCGGATGCGGAGGATCCCGAGAAAGCTCCAGTGTACCCCGAGGCGC 240

QY 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCGGAGTCTGGGTCCATC 300
DB 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCGGAGTCTGGGTCCATC 300

QY 301 ACCCTGCAAGTGCAGTCCGCCACCCAGGCGACCTTTCCTGCTCGGCTTTTAAGCAC 360
DB 301 ACCCTGCAAGTGCAGTCCGCCACCCAGGCGACCTTTCCTGCTCGGCTTTTAAGCAC 360

QY 361 AGCTCCCTGGGTGCCAGCGCGACTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 420
DB 361 AGCTCCCTGGGTGCCAGCGCGACTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 420

QY 421 ATCTTGAACGTGACAGACCCAGCGAGGAGATACCTTCTCATATTCAGACGCAACGC 480
DB 421 ATCTTGAACGTGACAGACCCAGCGAGGAGATACCTTCTCATATTCAGACGCAACGC 480

QY 481 GCCAACTACACAGTACTGTCACAGTGAATGTGAAGATACACAGCTGTATGCTGAAG 540
DB 481 GCCAACTACACAGTACTGTCACAGTGAATGTGAAGATACACAGCTGTATGCTGAAG 540

QY 541 AGACCTTACTTTAGGAAGATGGAACACCGAGATGCACCTCTCTGCTCTCCGAGGGTGT 600
DB 541 AGACCTTACTTTAGGAAGATGGAACACCGAGATGCACCTCTCTGCTCTCCGAGGGTGT 600

541 AGACCTTACTTTAGGAAGATGGAACACCGAGATGCACCTCTCTGCTCTCCGAGGGTGT 600
601 CCGGAGCCACATGTGGAGTGGTCTCTCAGCTCCACAGGAAAGCTGTAAAGAGAA 660
601 CCGGAGCCACATGTGGAGTGGTCTCTCAGCTCCACAGGAAAGCTGTAAAGAGAA 660
661 GCGCCTGTCTGTTGTCAGAAAGGAGGAAAGGTACTTTCATGAGTGTTCGGAACAGACATC 720
661 GCGCCTGTCTGTTGTCAGAAAGGAGGAAAGGTACTTTCATGAGTGTTCGGAACAGACATC 720
721 AGATGCTGTCTAGAAATGCACCTGGCGCGGATGCACCAAGCTGTTCACCATAGATCTA 780
721 AGATGCTGTCTAGAAATGCACCTGGCGCGGATGCACCAAGCTGTTCACCATAGATCTA 780
781 AACAGGCTCCTCAGAGCACACTGCCCCAGTTATCTTGAAGTGGGGAAACCTTGTGG 840
781 AACAGGCTCCTCAGAGCACACTGCCCCAGTTATCTTGAAGTGGGGAAACCTTGTGG 840
841 ATCAGGTGTAAAGCCCATCCATGTGAACCATGGATTCGGGCTCACCTGGGAGCTGGAAGAC 900
841 ATCAGGTGTAAAGCCCATCCATGTGAACCATGGATTCGGGCTCACCTGGGAGCTGGAAGAC 900
901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTTCCACAAACAGGACCATG 960
901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTTCCACAAACAGGACCATG 960
961 ATTGGGATTCCTTTGGCCCTTTGTTGCTCTCGTGGGAAAGAACGACACCGGATATTACAC 1020
961 ATTGGGATTCCTTTGGCCCTTTGTTGCTCTCGTGGGAAAGAACGACACCGGATATTACAC 1020
1021 TGCTCTTCTCAAAAGCACCCAGCCAGTCAGCGTTGGTGACCATCTCTAGAAAAGGGTTT 1080
1021 TGCTCTTCTCAAAAGCACCCAGCCAGTCAGCGTTGGTGACCATCTCTAGAAAAGGGTTT 1080
1081 ATAAAGCTACACAGCTCGCAAGAGAGTATGAAATTTGACCCGTACGAAAAGTTCTGCTTC 1140
1081 ATAAAGCTACACAGCTCGCAAGAGAGTATGAAATTTGACCCGTACGAAAAGTTCTGCTTC 1140
1141 TCAGTCAGGTTTAAAGCTACCCACGAATCCGATGCACGTGGATCTTCTCTCAAGCCTCA 1200
1141 TCAGTCAGGTTTAAAGCTACCCACGAATCCGATGCACGTGGATCTTCTCTCAAGCCTCA 1200
1201 TTTCTCTGTGAACAGAGAGGCGCTGGAGGTGGGTACAGCATATCTAAATTTTGGCATCAT 1260
1201 TTTCTCTGTGAACAGAGAGGCGCTGGAGGTGGGTACAGCATATCTAAATTTTGGCATCAT 1260
1261 AAGAACAAGCCAGGAGAGTACATATTTATGCAAGAAATGATGACGCCAGTTTCAACAAA 1320
1261 AAGAACAAGCCAGGAGAGTACATATTTATGCAAGAAATGATGACGCCAGTTTCAACAAA 1320
1321 ATGTTCCAGCTGAATATAGAAAGAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG 1380
1321 ATGTTCCAGCTGAATATAGAAAGAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG 1380
1381 GCCTCTGTCTCTCTGATGGCTACCCGCTACCCGCTACCCGCTGGAAGAGTGTTCGGAC 1440
1381 GCCTCTGTCTCTCTGATGGCTACCCGCTACCCGCTACCCGCTGGAAGAGTGTTCGGAC 1440
1441 AAATCTCCCAATTTGACGAGGAGAAATCCCAAGAGGAGTTTGGAAATAAAAGGCTTAACAGA 1500
1441 AAATCTCCCAATTTGACGAGGAGAAATCCCAAGAGGAGTTTGGAAATAAAAGGCTTAACAGA 1500
1501 AAAGTGTTCGCGAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGGAGCCGGAAGAGG 1560
1501 AAAGTGTTCGCGAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGGAGCCGGAAGAGG 1560
1561 CTCTCTGCTCAAAATGCTGTGCTCAATTTCTATGGCACGCTTTCGCAAAACCATCTTTT 1620
1561 CTCTCTGCTCAAAATGCTGTGCTCAATTTCTATGGCACGCTTTCGCAAAACCATCTTTT 1620
1621 AACTCACCAGGCGCCCTTCCCTTTTCATCCAAAGACAACTCTCCTTCTATGGACCATTTGGG 1680
1621 AACTCACCAGGCGCCCTTCCCTTTTCATCCAAAGACAACTCTCCTTCTATGGACCATTTGGG 1680
```



CLASSIFICATION: 536  
PRIOR APPLICATION DATA: US/07/813,593  
FILING DATE: 24-DEC-1991  
PRIOR APPLICATION DATA: US/07/793,065  
FILING DATE: 15-NOV-1991  
PRIOR APPLICATION DATA: US/07/728,913  
FILING DATE: 28-JUN-1991  
PRIOR APPLICATION DATA: US/07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-PPP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3453 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 31..3009  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 31..3006  
US-07-946-507-1

Query Match 100.0%; Score 3453; DB 1; Length 3453;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGGCGCTGCTACCGCGCGCTCCGGAGGCGCATCGGGCGTGGCGGAGCGCGACCGG 60  
DB 1 GGGCGCTGCTACCGCGCGCTCCGGAGGCGCATCGGGCGTGGCGGAGCGCGACCGG 60  
QY 61 CGGCTGCTGATCAAGTGTGTTTAAATCAGTCAATGATGAGACCGTTACAAACCAAG 120  
DB 61 CGGCTGCTGCTGCTGTTTGTTCAGTAAATGATGAGACCGTTACAAACCAAG 120  
QY 121 CTGCTGCTGATCAAGTGTGTTTAAATCAGTCAATGAGACCAATGGCTCATCAGCGG 180  
DB 121 CTGCTGCTGATCAAGTGTGTTTAAATCAGTCAATGAGACCAATGGCTCATCAGCGG 180  
QY 181 CCATCATGCTACCGAATGTCGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGCGCG 240  
DB 181 CCATCATGCTACCGAATGTCGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGCGCG 240  
QY 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGTGGCGGAGTCTGGGTCCATC 300  
DB 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGTGGCGGAGTCTGGGTCCATC 300  
QY 301 ACCCTGCAAGTGCAGCTGCCACCCAGGGGACCTTTCCCTGCTGCGTCTGGGTCTTTAAG 360  
DB 301 ACCCTGCAAGTGCAGCTGCCACCCAGGGGACCTTTCCCTGCTGCGTCTGGGTCTTTAAG 360  
QY 361 AGCTCCCTGGGCTGCCAGCGCGACTTTTGAATTTACAAACAGAGAAATCGTTCCATGGCC 420  
DB 361 AGCTCCCTGGGCTGCCAGCGCGACTTTTGAATTTACAAACAGAGAAATCGTTCCATGGCC 420  
QY 421 ATCTTGACCTGACAGAGACCCAGGAGAGATACCTACTCCATATTCAGAGCGAACGC 480  
DB 421 ATCTTGACCTGACAGAGACCCAGGAGAGATACCTACTCCATATTCAGAGCGAACGC 480  
QY 481 GCCAACTACACAGTACTGTTTCAGAGTGAATGTAAGAGATACACAGCTGTATGTCTAAG 540  
DB 481 GCCAACTACACAGTACTGTTTCAGAGTGAATGTAAGAGATACACAGCTGTATGTCTAAG 540

481 GCCAACTACACAGTACTGTTTCAGAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGG 540  
541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCACTGCTCTGCATCTCCGAGGGTGT 600  
541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCACTGCTCTGCATCTCCGAGGGTGT 600  
601 CCGGAGCCACTGTGGAGTGGGTGCTCTGCAGCTCCACAGGAAAGCTGTAAAGAGAA 660  
601 CCGGAGCCACTGTGGAGTGGGTGCTCTGCAGCTCCACAGGAAAGCTGTAAAGAGAA 660  
661 GGGCTGCTGTTGTCAGAAAGAGGAAAGTACTTTCATCAGTGTGTTTCGGAACAGACATC 720  
661 GGGCTGCTGTTGTCAGAAAGAGGAAAGTACTTTCATCAGTGTGTTTCGGAACAGACATC 720  
721 AGATGCTGTCTAGAAATGCACCTGGGCGCGAATGCACCAAGCTGTTTCCACATAGATCTA 780  
721 AGATGCTGTCTAGAAATGCACCTGGGCGCGAATGCACCAAGCTGTTTCCACATAGATCTA 780  
781 AACCAGGCTCTCAGAGCAGACTGCCCCAGTTATTCCTGAAGTGGGGGACCCCTTGTGG 840  
781 AACCAGGCTCTCAGAGCAGACTGCCCCAGTTATTCCTGAAGTGGGGGACCCCTTGTGG 840  
841 ATCAGGTGTAAGGCCATCCATGTGAACCATGATGCGGCTCAGCTGGAGCTGGAAGAC 900  
841 ATCAGGTGTAAGGCCATCCATGTGAACCATGATGCGGCTCAGCTGGAGCTGGAAGAC 900  
901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 960  
901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 960  
961 ATTCCGATTCCTTTGGGCTTTGTTGTTCCGTTGGGAAGGAGACGACCGGATATACACC 1020  
961 ATTCCGATTCCTTTGGGCTTTGTTGTTCCGTTGGGAAGGAGACGACCGGATATACACC 1020  
1021 TGCTCTTCTCAAGCACCCAGCTCAGCGTGGTGTGATGACCATCTCTAGAAAAAGGGTTT 1080  
1021 TGCTCTTCTCAAGCACCCAGCTCAGCGTGGTGTGATGACCATCTCTAGAAAAAGGGTTT 1080  
1081 ATAAAGCTACCAGCTCGCAAGAGAGTATGAAATTGACCCGTTACGAAAGTCTGCTTC 1140  
1081 ATAAAGCTACCAGCTCGCAAGAGAGTATGAAATTGACCCGTTACGAAAGTCTGCTTC 1140  
1141 TCAGTCAAGTTTAAAGCTACCCAGCAATCCGATGCGAGTGGATCTCTCTCAGACCTCA 1200  
1141 TCAGTCAAGTTTAAAGCTACCCAGCAATCCGATGCGAGTGGATCTCTCTCAGACCTCA 1200  
1201 TTTCCCTTGTCAACAGAGAGCGCTCGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
1201 TTTCCCTTGTCAACAGAGAGCGCTCGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
1261 AAGAAACAAGCCAGGAGAGTACATATTTCTATGCAAGAAATGATGACGCCAGTTCACCAA 1320  
1261 AAGAAACAAGCCAGGAGAGTACATATTTCTATGCAAGAAATGATGACGCCAGTTCACCAA 1320  
1321 ATGTTACAGCTGAATATAGAAAGAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG 1380  
1321 ATGTTACAGCTGAATATAGAAAGAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG 1380  
1381 GCGTCTGTTCTCTGATGCTACCCCTACCCCTCTTGGACCTGGAAGAGTGTTCGGAC 1440  
1381 GCGTCTGTTCTCTGATGCTACCCCTACCCCTCTTGGACCTGGAAGAGTGTTCGGAC 1440  
1441 AAATCTCCCAATTCGACGAGGAAATCCCAAGAGAGTGTGGAATAAAAAGCTAACAGA 1500  
1441 AAATCTCCCAATTCGACGAGGAAATCCCAAGAGAGTGTGGAATAAAAAGCTAACAGA 1500  
1501 AAAGTGTGTTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATGATGAGCGCGGGAAGGG 1560  
1501 AAAGTGTGTTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATGATGAGCGCGGGAAGGG 1560  
1561 CTTCGTGTCATGCTGTGCGTCAATCTTATGCGACGCTCTGCGAAACCATCTTTTAA 1620  
1561 CTTCGTGTCATGCTGTGCGTCAATCTTATGCGACGCTCTTGGGAAACCATCTTTTAA 1620



1621 AACTCACCAGGCCCTTCCCTTTTCATCCAAAGAACATCTCCTCTATCGACCACTTGG 1680  
1621 AACTCACCAGGCCCTTCCCTTTTCATCCAAAGAACATCTCCTCTATCGACCACTTGG 1680  
1681 CTCCTCTCCCTTCCCTTCTCATTTGTTGTTGATCTGCCAATAATACAAAAGCAA 1740  
1681 CTCCTCTCCCTTCCCTTCTCATTTGTTGTTGATCTGCCAATAATACAAAAGCAA 1740  
1741 TTATAGTACGAGAGCTCAGCTCAGATGATCCAGGTGACTGGGCCCTTGATACAGGTAC 1800  
1741 TTATAGTACGAGAGCTCAGCTCAGATGATCCAGGTGACTGGGCCCTTGATACAGGTAC 1800  
1801 TTCTACGTTGACTTCCAGGACTATCAATATGACCTTAAGTGGAGTTCCCGAGAGAGA 1860  
1801 TTCTACGTTGACTTCCAGGACTATCAATATGACCTTAAGTGGAGTTCCCGAGAGAGA 1860  
1861 TTAGAGTTTGGAGAGTCTCGGGTCTGGCGTTTCGGGAGGGTGAAGACCCAGGCC 1920  
1861 TTAGAGTTTGGAGAGTCTCGGGTCTGGCGTTTCGGGAGGGTGAAGACCCAGGCC 1920  
1921 TATGCACTTAAACGGAGTCTCAATTCAGGTGGCGTCAAGTGTCAAGTGTCAAGAGAAA 1980  
1921 TATGCACTTAAACGGAGTCTCAATTCAGGTGGCGTCAAGTGTCAAGTGTCAAGAGAAA 1980  
1981 GCTGACAGCTGTGAAAAGAGCTCTCATGTCGAGCTCAAAATGATGACCCACCTGGGA 2040  
1981 GCTGACAGCTGTGAAAAGAGCTCTCATGTCGAGCTCAAAATGATGACCCACCTGGGA 2040  
2041 CACCATGACACATCGTGAATCTGTGGGGGATGCACACTGTGAGGGCAGTGTACTTG 2100  
2041 CACCATGACACATCGTGAATCTGTGGGGGATGCACACTGTGAGGGCAGTGTACTTG 2100  
2101 ATTTTGAATATTTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 2160  
2101 ATTTTGAATATTTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 2160  
2161 TTTTACAGGACATGACAGAGATTTTAAAGAACATATTTTACAGTCTTACCTACTTTC 2220  
2161 TTTTACAGGACATGACAGAGATTTTAAAGAACATATTTTACAGTCTTACCTACTTTC 2220  
2221 CAGGACATTCAAATTCACGATGCTGTTTACGAGAGTTTACAGTCTTACCTACTTTC 2280  
2221 CAGGACATTCAAATTCACGATGCTGTTTACGAGAGTTTACAGTCTTACCTACTTTC 2280  
2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTCATTCATTCATTCATTCATTCATTCATTC 2340  
2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTCATTCATTCATTCATTCATTCATTCATTC 2340  
2341 AACCAAGAGAGCTGCGAG 2400  
2341 AACCAAGAGAGCTGCGAG 2400  
2401 CTCCTTTGCTTGTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGCT 2460  
2401 CTCCTTTGCTTGTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGCT 2460  
2461 GTCCACAGAGACCTGGCAGCAGGAATGTTGTTGTCACCCAGGAGGTGTTGAGATGAA 2520  
2461 GTCCACAGAGACCTGGCAGCAGGAATGTTGTTGTCACCCAGGAGGTGTTGAGATGAA 2520  
2521 TGTGACTTGGACTGGCCGAGACATCTGAGGAGTCCAGCTACCTGCTGAGGGCAAC 2580  
2521 TGTGACTTGGACTGGCCGAGACATCTGAGGAGTCCAGCTACCTGCTGAGGGCAAC 2580  
2581 GCACGGCTGCCGTTGAAGTGGATGGCACCAGAGAGCTTATTTGAAGGATCTACACAATC 2640  
2581 GCACGGCTGCCGTTGAAGTGGATGGCACCAGAGAGCTTATTTGAAGGATCTACACAATC 2640  
2641 AAGAGTACGCTGTGCTTACGGCATCTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700  
2641 AAGAGTACGCTGTGCTTACGGCATCTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700

2701 CCTTACCCTGGCATTCCTGTGCGAGCTAACTTCTATAAACTGATTTCAGAGTGGATTTAAA 2760  
2701 CCTTACCCTGGCATTCCTGTGCGAGCTAACTTCTATAAACTGATTTCAGAGTGGATTTAAA 2760  
2761 ATGGAGCAGCATTCATGCCCACAGAGGATATCTTTTAATGCAATCCTGCTGGCT 2820  
2761 ATGGAGCAGCATTCATGCCCACAGAGGATATCTTTTAATGCAATCCTGCTGGCT 2820  
2821 TTTGACTCAAGGAGGCGCATCTTCCCAACTGACTTCTATTTTAGGATGTCAGCTG 2880  
2821 TTTGACTCAAGGAGGCGCATCTTCCCAACTGACTTCTATTTTAGGATGTCAGCTG 2880  
2881 GCAGAGCAGAGCAATGATATCAGAACTCCATCCATACCAAAACAGGCGGCCCT 2940  
2881 GCAGAGCAGAGCAATGATATCAGAACTCCATCCATACCAAAACAGGCGGCCCT 2940  
2941 CAGCAGAGAGGCGGCTCAGAGCCAGTCCGACAGCGCCAGGTTGAAGATTCACAGAGAA 3000  
2941 CAGCAGAGAGGCGGCTCAGAGCCAGTCCGACAGCGCCAGGTTGAAGATTCACAGAGAA 3000  
3001 AGAAGTTAGGAGGCGCTTGGACCCCGCCCTAGCAGGCTGTAGCCGACAGAGCA 3060  
3001 AGAAGTTAGGAGGCGCTTGGACCCCGCCCTAGCAGGCTGTAGCCGACAGAGCA 3060  
3061 AGATTAGCTCGCTCTGAGAGAGCGCTTACAGCGCTTCTGCTGAGCTTTCTCT 3120  
3061 AGATTAGCTCGCTCTGAGAGAGCGCTTACAGCGCTTCTGCTGAGCTTTCTCT 3120  
3121 AGATGCTGTGCTTACTTCCAAAGTACTTCTATAAACTCAAACTCTCTCGCACAG 3180  
3121 AGATGCTGTGCTTACTTCCAAAGTACTTCTATAAACTCAAACTCTCTCGCACAG 3180  
3181 GCGGAGAGCCATATGAGACTTGTGTTGAGCGCCCTACCTGGGGGCTTTCCACG 3240  
3181 GCGGAGAGCCATATGAGACTTGTGTTGAGCGCCCTACCTGGGGGCTTTCCACG 3240  
3241 AGCTTACAGGAG 3300  
3241 AGCTTACAGGAG 3300  
3301 ACCAAACCCGTTTTTGTGCTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
3301 ACCAAACCCGTTTTTGTGCTAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360  
3361 AATACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420  
3361 AATACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420  
3421 TACTGTAAAAAAG 3453  
3421 TACTGTAAAAAAG 3453

## RESULT 4

US-08-252-517-1  
; Sequence 1, Application US/08252517  
; Patent No. 5548065  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS



Db 1201 TTTCTCTGTGAACAGAGAGGCGCTGGAGGATGGGTACAGCATATCTAAATTTTTCGGATCAT 1260  
QY 1261 AAGAACAAGCCAGGAGGTACATATTTCTATGAGAAAATGATGACGCCAGTTCCACCAA 1320  
Db 1261 AAGAACAAGCCAGGAGGTACATATTTCTATGAGAAAATGATGACGCCAGTTCCACCAA 1320  
QY 1321 ATGTTACAGCTGAATATATAAGAAAGAAACCTCAAGTGTCTAGCAAAATGCCCTCAGCCAG 1380  
Db 1321 ATGTTACAGCTGAATATATAAGAAAGAAACCTCAAGTGTCTAGCAAAATGCCCTCAGCCAG 1380  
QY 1381 GGTCTCTGTCTCTGATGGCTACCGGCTACCGCTACCGCTTGGACCTGGAAGAGTGTTCGGAC 1440  
Db 1381 GGTCTCTGTCTCTGATGGCTACCGGCTACCGCTACCGCTTGGACCTGGAAGAGTGTTCGGAC 1440  
QY 1441 AAATCTCCCAATGTCAGGAGGAAATCCAGAGGAGTTTGGAAATAAAAGGCTACACGA 1500  
Db 1441 AAATCTCCCAATGTCAGGAGGAAATCCAGAGGAGTTTGGAAATAAAAGGCTACACGA 1500  
QY 1501 AAAGTGTGTTGGCCAGTGGGTGTGAGCAGTACTCTAAATATAGTGAAGCCGGGAAGGG 1560  
Db 1501 AAAGTGTGTTGGCCAGTGGGTGTGAGCAGTACTCTAAATATAGTGAAGCCGGGAAGGG 1560  
QY 1561 CTTCTGGTCAATGCTGCGGTACAAATTTCTATGCGCACGCTCTGGCAAAACCATCTTTTAA 1620  
Db 1561 CTTCTGGTCAATGCTGCGGTACAAATTTCTATGCGCACGCTCTGGCAAAACCATCTTTTAA 1620  
QY 1621 AACTCAGCAGGCCCTTCCCTTTCATCCAGAGACATCTCTCTATCGGCCATTTGG 1680  
Db 1621 AACTCAGCAGGCCCTTCCCTTTCATCCAGAGACATCTCTCTATCGGCCATTTGG 1680  
QY 1681 CTCTCTCCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740  
Db 1681 CTCTCTCCTCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740  
QY 1741 TTTAGGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCCCTGGATACAGGTAC 1800  
Db 1741 TTTAGGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCCCTGGATACAGGTAC 1800  
QY 1801 TTCTACGTTGACTTCAGGAGTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAAAC 1860  
Db 1801 TTCTACGTTGACTTCAGGAGTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAAAC 1860  
QY 1861 TTAGAGTTTGGAGGTCCTGGGCTCTGGCGCTTTTCGGAGGGTGATGAACGCCAGGCC 1920  
Db 1861 TTAGAGTTTGGAGGTCCTGGGCTCTGGCGCTTTTCGGAGGGTGATGAACGCCAGGCC 1920  
QY 1921 TATGCAATTAGTAAACGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAA 1980  
Db 1921 TATGCAATTAGTAAACGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAA 1980  
QY 1981 GCTGACAGCTGTGAAGAAAGCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGGA 2040  
Db 1981 GCTGACAGCTGTGAAGAAAGCTCTCATGTGGAGCTCAAAATGATGACCCACCTGGGA 2040  
QY 2041 CACCATGACAACATCGTGAATCTGCTGGGGGATGCACACTGTGAGGGCCAGTGTACTTG 2100  
Db 2041 CACCATGACAACATCGTGAATCTGCTGGGGGATGCACACTGTGAGGGCCAGTGTACTTG 2100  
QY 2101 ATTTTGAATATGTTGCTATGGTGACCTCCAACTACCTAAGAGTAAAGAGAGAAG 2160  
Db 2101 ATTTTGAATATGTTGCTATGGTGACCTCCCACTACCTAAGAGTAAAGAGAGAAG 2160  
QY 2161 TTTACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCAGTCTTACCCCTACTTTC 2220  
Db 2161 TTTACAGGACATGGACAGAGATTTTAAAGAACATAATTTTCAGTCTTACCCCTACTTTC 2220  
QY 2221 CAGGCACATTCAAATCCAGCATGCTGGTTCCAGAGAGTTTCAGTTACACCCGCCCTTG 2280  
Db 2221 CAGGCACATTCAAATCCAGCATGCTGGTTCCAGAGAGTTTCAGTTACACCCGCCCTTG 2280  
QY 2281 GATCAGCTCTCAGGGTCAATGGGAATTCATTTCTGAAAGTGAAGTGAATATGAA 2340  
Db 2281 GATCAGCTCTCAGGGTCAATGGGAATTCATTTCTGAAAGTGAAGTGAATATGAA 2340

QY 2341 AACCAAGAAGGCTGGCAGAAAGAGAGGAGGAAAGATTGTAACGTGCTGACGTTTGAAGAC 2400  
Db 2341 AACCAAGAAGGCTGGCAGAAAGAGAGGAGGAAAGATTGTAACGTGCTGACGTTTGAAGAC 2400  
QY 2401 CTCCTTTGCTTTCCGTACCAAGTGGCCAAAGGATGAAATTCCTGGAGTTCAAGTCGTGT 2460  
Db 2401 CTCCTTTGCTTTCCGTACCAAGTGGCCAAAGGATGAAATTCCTGGAGTTCAAGTCGTGT 2460  
QY 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTTGGTCAACCCACGCGGAAGGTGGTCAAGATC 2520  
Db 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTTGGTCAACCCACGCGGAAGGTGGTCAAGATC 2520  
QY 2521 TGTGACTTTGGACTGGCCCGAGACATCTCAGCGACTCCAGCTACGTCGTCAAGGGCAAC 2580  
Db 2521 TGTGACTTTGGACTGGCCCGAGACATCTCAGCGACTCCAGCTACGTCGTCAAGGGCAAC 2580  
QY 2581 GCACGGCTCCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACACAATC 2640  
Db 2581 GCACGGCTCCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACACAATC 2640  
QY 2641 AAGAGTGACGTCTGTCTACGGCATCTCTCTGGAGATATTTTTCACGTGGGTGTGAAC 2700  
Db 2641 AAGAGTGACGTCTGTCTACGGCATCTCTCTGGAGATATTTTTCACGTGGGTGTGAAC 2700  
QY 2701 CCTTACCTGGCATCTCTGTGACGCTAACTTCTATAAAGTATTCAGAGTGGATTTAAA 2760  
Db 2701 CCTTACCTGGCATCTCTGTGACGCTAACTTCTATAAAGTATTCAGAGTGGATTTAAA 2760  
QY 2761 ATGAGCAGCATTCTATGCCACAGAGGATATACCTTTGTAATGCAATCTCTGCTGGCT 2820  
Db 2761 ATGAGCAGCATTCTATGCCACAGAGGATATACCTTTGTAATGCAATCTCTGCTGGCT 2820  
QY 2821 TTTGACTCAAGGAAGCGGCCATCTTCCCAACCTGACTTCAATTTTAGGATGTACGCTG 2880  
Db 2821 TTTGACTCAAGGAAGCGGCCATCTTCCCAACCTGACTTCAATTTTAGGATGTACGCTG 2880  
QY 2881 GCAGAGCAGAGAAGCATGTATCAGAACATCCATCCATACCAAAACAGGCGGCCCT 2940  
Db 2881 GCAGAGCAGAGAAGCATGTATCAGAACATCCATCCATACCAAAACAGGCGGCCCT 2940  
QY 2941 CAGCAGAGAGCGGGCTCAGAGCCAGCCAGCCAGCGGAGTGAAGATTACAGAGAA 3000  
Db 2941 CAGCAGAGAGCGGGCTCAGAGCCAGCCAGCCAGCGGAGTGAAGATTACAGAGAA 3000  
QY 3001 AGAAGTTAGGAGAGCGCTTTGGACCCCGCCCTAGCAGCGTGTAGACCGCAGAGCCA 3060  
Db 3001 AGAAGTTAGCAGAGCGCTTTGGACCCCGCCCTAGCAGCGTGTAGACCGCAGAGCCA 3060  
QY 3061 AGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCGCTGGACTTTTCTCT 3120  
Db 3061 AGATTAGCCTCGCTCTGAGGAAGCGCCCTACAGCGGTTGCTTCGCTGGACTTTTCTCT 3120  
QY 3121 AGATGCTGTCTGCTTACTTCCAAAGTACTTCTATAAAATCAAACTCTCTCCACAG 3180  
Db 3121 AGATGCTGTCTGCTTACTTCCAAAGTACTTCTATAAAATCAAACTCTCTCCACAG 3180  
QY 3181 GCGGAGAGCCAAATATAGACTTTTGGTGAGCGCCCTACCCCTGGGGGCTTTTTCAGC 3240  
Db 3181 GCGGAGAGCCAAATATAGACTTTTGGTGAGCGCCCTACCCCTGGGGGCTTTTTCAGC 3240  
QY 3241 AGCTTCAGGGGAAGCCATGTATCTGAATATAGTATATTTTGAATACGTGAACAA 3300  
Db 3241 AGCTTCAGGGGAAGCCATGTATCTGAATATAGTATATTTTGAATACGTGAACAA 3300  
QY 3301 ACCAAACCCGTTTTTTCGCTAAGGGAAGCTAAATATGATTTTAAAAATCTATGTTTTAA 3360  
Db 3301 ACCAAACCCGTTTTTTCGCTAAGGGAAGCTAAATATGATTTTAAAAATCTATGTTTTAA 3360  
QY 3361 AATACTATGTAATTTTTCATCTATTTAGTGATATATTTTATGATGGAATAAATTTTC 3420  
Db 3361 AATACTATGTAATTTTTCATCTATTTAGTGATATATTTTATGATGGAATAAATTTTC 3420

QY 3421 TACTGTAAAAA... 3453  
Db 3421 TACTGTAAAAA... 3453

RESULT 5  
US-07-906-397A-1  
Sequence 1, Application US/07906397A  
Patent No. 5621090

GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: IMCLONE SYSTEMS INCORPORATED  
STREET: 180 VARICK STREET  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10014

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/906,397A  
FILING DATE: 19920626  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3453 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 31..3009  
NAME/KEY: mat\_peptide  
LOCATION: 31..3006  
US-07-906-397A-1

Query Match 100.0%; Score 3453; DB 1; Length 3453;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCTGGCTACCGCGGCTCCGGAGGCCATCGCGGGCTTGGCGCAGCGCAGCGACCGG 60  
Db 1 CGCGCTGGCTACCGCGGCTCCGGAGGCCATCGCGGGCTTGGCGCAGCGCAGCGACCGG 60

1 CGCGCTGGCTACCGCGGCTCCGGAGGCCATCGCGGGCTTGGCGCAGCGCAGCGACCGG 60  
61 CGCGCTGGCTACCGCGGCTCCGGAGGCCATCGCGGGCTTGGCGCAGCGCAGCGACCGG 120  
61 CGCGCTGGCTACCGCGGCTCCGGAGGCCATCGCGGGCTTGGCGCAGCGCAGCGACCGG 120  
121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCAATGAGACAATGGCTCATCAGCGGGAAG 180  
121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCAATGAGACAATGGCTCATCAGCGGGAAG 180  
181 CCATCATCGTACCGAATGGTGGAGATCCCGAGAGACCTCCAGTGTACCCGAGCGGC 240  
181 CCATCATCGTACCGAATGGTGGAGATCCCGAGAGACCTCCAGTGTACCCGAGCGGC 240  
241 CAGAGTGAAGGACGATATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGTCCATC 300  
241 CAGAGTGAAGGACGATATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGTCCATC 300  
301 ACCCTGCAAGTGCAGTCCCGACCGGACCTTTCTGCTCTGGGTCTTTAAGCAC 360  
301 ACCCTGCAAGTGCAGTCCCGACCGGACCTTTCTGCTCTGGGTCTTTAAGCAC 360  
361 AGCTCCCTGGGCTGCCAGCGGCACCTTTGATTTACAAAACAGAGAAATCGTTCCATGGCC 420  
361 AGCTCCCTGGGCTGCCAGCGGCACCTTTGATTTACAAAACAGAGAAATCGTTCCATGGCC 420  
421 ATCTTGAAGTGCAGAGACCGGAGGAGATACCTTCTCATATTCAGAGCGAACGC 480  
421 ATCTTGAAGTGCAGAGACCGGAGGAGATACCTTCTCATATTCAGAGCGAACGC 480  
481 GCCAACTACACAGTACTGTTACAGTGAATGTAAAGATACACAGCTGTATGTCTAAGG 540  
481 GCCAACTACACAGTACTGTTACAGTGAATGTAAAGATACACAGCTGTATGTCTAAGG 540  
541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCTCTCTGCTCTCCGAGGGTGT 600  
541 AGACCTTACTTTAGGAAGATGGAACACAGGATGCTCTCTGCTCTCCGAGGGTGT 600  
601 CGGAGCGCACCTGGAGTGGGTGCTCTGAGCTCCACAGGAGAAAGCTGTAAGAAAGAA 660  
601 CGGAGCGCACCTGGAGTGGGTGCTCTGAGCTCCACAGGAGAAAGCTGTAAGAAAGAA 660  
661 GGCCCTGCTGTTGTCAGAAAGAGGAAAGGTACTTCTATGAGTGTTCGGAACAGACATC 720  
661 GGCCCTGCTGTTGTCAGAAAGAGGAAAGGTACTTCTATGAGTGTTCGGAACAGACATC 720  
721 AGATGCTGTGCTAGAAATGCAGTGGCGCGGAATGCACCAAGCTGTTTACCATAGATCTA 780  
721 AGATGCTGTGCTAGAAATGCAGTGGCGCGGAATGCACCAAGCTGTTTACCATAGATCTA 780  
781 AACGAGCTCCTCAGAGCACCTGCCCGAGTATTTCTGAAAGTGGGGGAAACCCCTTGG 840  
781 AACGAGCTCCTCAGAGCACCTGCCCGAGTATTTCTGAAAGTGGGGGAAACCCCTTGG 840  
841 ATCAGGTGTAAGGCCCATCCATGTGAACCATGGATTTCGGGCTCACTGGAGCTGGAAGAC 900  
841 ATCAGGTGTAAGGCCCATCCATGTGAACCATGGATTTCGGGCTCACTGGAGCTGGAAGAC 900  
901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCCAACAGACACCATG 960  
901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCCAACAGACACCATG 960  
961 ATTCGGATTTCTTGGCCTTTGCTCTCCGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
961 ATTCGGATTTCTTGGCCTTTGCTCTCCGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
1021 TGCTCTTCTCAAAAGCACCCAGCGTGTGGTGTGAGCATCTCTAGAAAAGAGGGTTT 1080  
1021 TGCTCTTCTCAAAAGCACCCAGCGTGTGGTGTGAGCATCTCTAGAAAAGAGGGTTT 1080  
1081 ATAAACGCTACCGCTCGCAAGAGAGATGAAATTGACCCGCTACGAAAAGTCTGCTTC 1140  
1081 ATAAACGCTACCGCTCGCAAGAGAGATGAAATTGACCCGCTACGAAAAGTCTGCTTC 1140

1141 TCAGTCAGGTTTAAAGCGTACCCAGAAATCCGATGCACGTGATCTTCTCTCAAGCCTCA 1200  
1141 TCAGTCAGGTTTAAAGCGTACCCAGAAATCCGATGCACGTGATCTTCTCTCAAGCCTCA 1200  
1201 TTTCCCTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTTCGATCAT 1260  
1201 TTTCCCTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTTCGATCAT 1260  
1261 AAGACAGCCAGGAGAGTACATATCTATGACAGAAATGATGACGCCAGTTCACCAAA 1320  
1261 AAGACAGCCAGGAGAGTACATATCTATGACAGAAATGATGACGCCAGTTCACCAAA 1320  
1321 ATGTTTCAGGCTGAATATAGAAGAAACCTCAAGTCTAGCAAAATGCCCTCAGCCAGCCAG 1380  
1321 ATGTTTCAGGCTGAATATAGAAGAAACCTCAAGTCTAGCAAAATGCCCTCAGCCAGCCAG 1380  
1381 GCGTCCCTGTCTCTGATGGCTACCCGCTACCCCTCTTGACCTTGGAAAGTGTTCGGAC 1440  
1381 GCGTCCCTGTCTCTGATGGCTACCCGCTACCCCTCTTGACCTTGGAAAGTGTTCGGAC 1440  
1441 AAATCTCCCAATTCACGAGGAAATCCAGAGGAGTTTGGAAATAGAAAGGCTACACAA 1500  
1441 AAATCTCCCAATTCACGAGGAAATCCAGAGGAGTTTGGAAATAGAAAGGCTACACAA 1500  
1501 AAAGTGTGTGGCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGAGGCCGGGAAAGG 1560  
1501 AAAGTGTGTGGCAGTGGGTGTCGAGCAGTACTCTAAATATGAGTGAGGCCGGGAAAGG 1560  
1561 CTTCTGGTCAATGCTGTGCGTACAAATCTATGGCACGCTTGGGAAACCATCTTTTAA 1620  
1561 CTTCTGGTCAATGCTGTGCGTACAAATCTATGGCACGCTTGGGAAACCATCTTTTAA 1620  
1621 AACTCAGAGGCCCTTCCCTTTCATCCAGACACATCTCTTCTATCGGACCATTTGG 1680  
1621 AACTCAGAGGCCCTTCCCTTTCATCCAGACACATCTCTTCTATCGGACCATTTGG 1680  
1681 CTCTGTCTCCCTTCATTTGTTCTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740  
1681 CTCTGTCTCCCTTCATTTGTTCTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740  
1741 TTTAGTACGAGAGTCAGCTGCAGATGATCCAGTGACTGGCCCCCTGGATACAGAGTAC 1800  
1741 TTTAGTACGAGAGTCAGCTGCAGATGATCCAGTGACTGGCCCCCTGGATACAGAGTAC 1800  
1801 TTTCTAGTGTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCTCGAGAGAA 1860  
1801 TTTCTAGTGTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCTCGAGAGAA 1860  
1861 TTAGAGTTTGGAAAGTCTCTGGGTCTGGCGCTTTTGGGAGGTTGATGAACGCCAGGCC 1920  
1861 TTAGAGTTTGGAAAGTCTCTGGGTCTGGCGCTTTTGGGAGGTTGATGAACGCCAGGCC 1920  
1921 TATGCAATTTAGTAAACGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA 1980  
1921 TATGCAATTTAGTAAACGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA 1980  
1981 GCTGACAGCTGTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTTGGGA 2040  
1981 GCTGACAGCTGTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTTGGGA 2040  
2041 CACCATGACAACTCGTGAATCTGTGGGGGATGCACACTGTGACGGCCAGTGTACTTGG 2100  
2041 CACCATGACAACTCGTGAATCTGTGGGGGATGCACACTGTGACGGCCAGTGTACTTGG 2100  
2101 ATTTTGTGAATATGTTGCTATGTTGACCTCCCTCACTACCTAAGAAATGAAAGAGAGAAG 2160  
2101 ATTTTGTGAATATGTTGCTATGTTGACCTCCCTCACTACCTAAGAAATGAAAGAGAGAAG 2160  
2161 TTTTCAGGACATGACAGAGATTTTAAAGGAACATAATTTTCAGTTCTTACCCCTACTTTC 2220  
2161 TTTTCAGGACATGACAGAGATTTTAAAGGAACATAATTTTCAGTTCTTACCCCTACTTTC 2220

2221 CAGGCACATTTCAAATTCACGATGCCTGGTTACGAGAAAGTTTCAGTTACACCGCCCTTG 2280  
2221 CAGGCACATTTCAAATTCACGATGCCTGGTTACGAGAAAGTTTCAGTTACACCGCCCTTG 2280  
2281 GATCAGCTCTCAGGTTTCAATGGGAATTCAAATTCATTTCTCAAGATGAGATTGAATGAA 2340  
2281 GATCAGCTCTCAGGTTTCAATGGGAATTCAAATTCATTTCTCAAGATGAGATTGAATGAA 2340  
2341 ACCCAGAGAGCTGCCAGAGAGAGGAGGAATTTGAACGTGCTGAGCTTTGAAGAC 2400  
2341 ACCCAGAGAGCTGCCAGAGAGAGGAGGAATTTGAACGTGCTGAGCTTTGAAGAC 2400  
2401 CTCCTTTGCTTTGCGTACCAAGTGGCCAAAGGATGAAATTTCTTGGAGTTCAAGTCGTG 2460  
2401 CTCCTTTGCTTTGCGTACCAAGTGGCCAAAGGATGAAATTTCTTGGAGTTCAAGTCGTG 2460  
2461 GTCCACAGAGACTGGCAGCCAGGAATGTGTTGGTCAACCCAGGGAAGGTGGTGAAGATC 2520  
2461 GTCCACAGAGACTGGCAGCCAGGAATGTGTTGGTCAACCCAGGGAAGGTGGTGAAGATC 2520  
2521 TGTGACTTTGGACTGGCCGAGACATCTCTCAGCGACTCCAGCTACGTCAGCGGGCAAC 2580  
2521 TGTGACTTTGGACTGGCCGAGACATCTCTCAGCGACTCCAGCTACGTCAGCGGGCAAC 2580  
2581 GCACGGCTCCGCTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACACAATC 2640  
2581 GCACGGCTCCGCTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACACAATC 2640  
2641 AAGAGTACGCTGTGCTTACGGCATCTTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700  
2641 AAGAGTACGCTGTGCTTACGGCATCTTCTCTGGAGATATTTTCACTGGGTGTGAAC 2700  
2701 CCTTACCTTGGCATCTCTGTCGACGCTAACTTCTATAAAGTATTCAGAGTGGATTTAA 2760  
2701 CCTTACCTTGGCATCTCTGTCGACGCTAACTTCTATAAAGTATTCAGAGTGGATTTAA 2760  
2761 ATGGAGCAGCATCTATGCCACAGAGGATATCTTCTTAATGCAATCTCTGCTGGCT 2820  
2761 ATGGAGCAGCATCTATGCCACAGAGGATATCTTCTTAATGCAATCTCTGCTGGCT 2820  
2821 TTTGACTCAAGGAGCGGCCATCTTCCCCAACCTGACTTCAATTTTGGATGTGAGCTG 2880  
2821 TTTGACTCAAGGAGCGGCCATCTTCCCCAACCTGACTTCAATTTTGGATGTGAGCTG 2880  
2881 GCAGAGCAGAGAAAGCATGTATCAGAACATCCATCCATACCAAAACAGGCGGCCCT 2940  
2881 GCAGAGCAGAGAAAGCATGTATCAGAACATCCATCCATACCAAAACAGGCGGCCCT 2940  
2941 CAGCAGAGAGCGGCTCAGAGCCCGCAGCGCCAGCGGCTGAAGATTTCACAGAGAA 3000  
2941 CAGCAGAGAGCGGCTCAGAGCCCGCAGCGCCAGCGGCTGAAGATTTCACAGAGAA 3000  
3001 AGAAGTTAGCGAGGCGCTTGGACCCCGCCCTAGCAGGCTGTAGACCGCAGAGCCA 3060  
3001 AGAAGTTAGCGAGGCGCTTGGACCCCGCCCTAGCAGGCTGTAGACCGCAGAGCCA 3060  
3061 AGATTAGCTCGCTCTGAGGAAGCGCCCTACAGCGGCTTCTGCTGACACTTTTCTCT 3120  
3061 AGATTAGCTCGCTCTGAGGAAGCGCCCTACAGCGGCTTCTGCTGACACTTTTCTCT 3120  
3121 AGATTGCTGTGCTGCTACTTCCAAAGTACTTCTATAAATCAAACTCTCTCGCACAG 3180  
3121 AGATTGCTGTGCTGCTACTTCCAAAGTACTTCTATAAATCAAACTCTCTCGCACAG 3180  
3181 GCGGAGAGCCAAATAGACTTTGTTGGTGAAGCGCCCTACCTTGGGCGCTTTTCCACG 3240  
3181 GCGGAGAGCCAAATAGACTTTGTTGGTGAAGCGCCCTACCTTGGGCGCTTTTCCACG 3240  
3241 AGCTTGAGGGAAGCCATGTATCTGAATATAGTATATTTCTGTAATACGTGAACAA 3300  
3241 AGCTTGAGGGAAGCCATGTATCTGAATATAGTATATTTCTGTAATACGTGAACAA 3300  
3301 ACCAAACCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTTTAAAAATCTATGTTTTAA 3360

Db 3301 ACCAACCCGTTTTCGTAAGGAAAGCTAAATATGATTTTAAAAATCTATGTTTAA 3360  
QY 3361 AATACATGTAACCTTTTCATCTATTTAGTGATATATTTATGGATGGAATAAACTTTC 3420  
Db 3361 AATACATGTAACCTTTTCATCTATTTAGTGATATATTTATGGATGGAATAAACTTTC 3420  
QY 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453  
Db 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453

RESULT 6  
US-08-601-891-1  
Sequence 1, Application US/08601891  
Patent No. 5747651  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ImClone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/601,891  
FILING DATE: 15-FEB-1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/977,451  
FILING DATE: 19-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:

LENGTH: 3453 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 112..3006  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 31..111  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 31..3009  
US-08-601-891-1  
Query Match 100.0%; Score 3453; DB 1; Length 3453;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCGGCCTGGCTACCGCGCGCTCCGAGGCGCATCGGGCGTGGCGCAGCGCAGCGCGG 60  
Db 1 GCGGCCTGGCTACCGCGCGCTCCGAGGCGCATCGGGCGTGGCGCAGCGCAGCGCGG 60  
QY 61 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
Db 61 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
QY 121 CTGCGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGCTCATCAGCGGGAAG 180  
Db 121 CTGCGCTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGCTCATCAGCGGGAAG 180  
QY 181 CCATCATCGTACCGAATGTTGCGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGCGC 240  
Db 181 CCATCATCGTACCGAATGTTGCGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGCGC 240  
QY 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC 300  
Db 241 CAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC 300  
QY 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGAGCTTTCCTGCGCTCTGGGTCTTTAAGCAC 360  
Db 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGGGAGCTTTCCTGCGCTCTGGGTCTTTAAGCAC 360  
QY 361 AGCTCCCTGGGCTGCCAGCGCCACTTTGATTTACAAAACAGAGGAATCGTTTCCATGCC 420  
Db 361 AGCTCCCTGGGCTGCCAGCGCCACTTTGATTTACAAAACAGAGGAATCGTTTCCATGCC 420  
QY 421 ATCTTGAAGCTGACAGAGACCCAGGAGGAGATACCTACTCCATATTTACAGAGCGACGC 480  
Db 421 ATCTTGAAGCTGACAGAGACCCAGGAGGAGATACCTACTCCATATTTACAGAGCGACGC 480  
QY 481 GCCAACTACACAGTACTGTTACAGTGAATGTAGAGATACAGCTGTATGTCTAGG 540  
Db 481 GCCAACTACACAGTACTGTTACAGTGAATGTAGAGATACAGCTGTATGTCTAGG 540  
QY 541 AGACCTTACTTTAGGAAGATGGAACCCAGGATGCACTGCTGCTCATCTCCGAGGGTGT 600  
Db 541 AGACCTTACTTTAGGAAGATGGAACCCAGGATGCACTGCTGCTCATCTCCGAGGGTGT 600  
QY 601 CCGAGCGCCACTGTGGAGTGGGTGCTGTCAGTCCACAGGAAAGCTGTAAGAGAA 660  
Db 601 CCGAGCGCCACTGTGGAGTGGGTGCTGTCAGTCCACAGGAAAGCTGTAAGAGAA 660  
QY 661 GGCCTGCTGTTGTCAGAAAGGAGGAAAGGTTACTTATCATGTTGTTGCGAAGACATC 720  
Db 661 GGCCTGCTGTTGTCAGAAAGGAGGAAAGGTTACTTATCATGTTGTTGCGAAGACATC 720  
QY 721 AGATGCTGTGCTAGAAATGCACCTGGCGCGGAATGCACCAAGCTGTTTCCATAGATCTA 780  
Db 721 AGATGCTGTGCTAGAAATGCACCTGGCGCGGAATGCACCAAGCTGTTTCCATAGATCTA 780





QY 2941 CAGCAGAGAGGGGGCTCAGAGCCAGTCCACAGCGCCAGTGAAGATTACAGAGAA 3000  
Db 2941 CAGCAGAGAGGGGGCTCAGAGCCAGTCCACAGCGCCAGTGAAGATTACAGAGAA 3000  
QY 3001 AGAAGTTAGCGAGAGGGCTTGAGCCCGCCAGCTGAGAGGCTGAGAGCCGAGAGCCA 3060  
Db 3001 AGAAGTTAGCGAGAGGGCTTGAGCCCGCCAGCTGAGAGGCTGAGAGCCGAGAGCCA 3060  
QY 3061 AGATTAGCTCCCTCTGAGGAGCGCCCTACAGCGCGTTCCTCGACTTTTCTCT 3120  
Db 3061 AGATTAGCTCCCTCTGAGGAGCGCCCTACAGCGCGTTCCTCGACTTTTCTCT 3120  
QY 3121 AGATGCTGTGCTGCAATTAATCAAGTGAATCTTATAAATCAAACTCTCTCGCACAG 3180  
Db 3121 AGATGCTGTGCTGCAATTAATCAAGTGAATCTTATAAATCAAACTCTCTCGCACAG 3180  
QY 3181 GCGGAGAGCCAAATAGACTTTGTTGGTGAGCCCGCTACCCCTGGGGCCCTTTCCACG 3240  
Db 3181 GCGGAGAGCCAAATAGACTTTGTTGGTGAGCCCGCTACCCCTGGGGCCCTTTCCACG 3240  
QY 3241 AGCTTGAGGGAAGCCATCTATCTGAATATAGTATATCTTGTAAATACGTGAACAA 3300  
Db 3241 AGCTTGAGGGAAGCCATCTATCTGAATATAGTATATCTTGTAAATACGTGAACAA 3300  
QY 3301 ACCAAACCGCTTTTGTCTAAGGGAAGCTAAATATGATTTTAAAAATCTATGTTTAA 3360  
Db 3301 ACCAAACCGCTTTTGTCTAAGGGAAGCTAAATATGATTTTAAAAATCTATGTTTAA 3360  
QY 3361 ATACTATGACTTTTTCATCTATTAGTATATTTATGATATTTATGATGGAATAAATTC 3420  
Db 3361 ATACTATGACTTTTTCATCTATTAGTATATTTATGATATTTATGATGGAATAAATTC 3420  
QY 3421 TACTGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3453  
Db 3421 TACTGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3453

## RESULT 7

US-09-021-324-1

Sequence 1, Application US/09021324

Patent No. 5912133

GENERAL INFORMATION:

APPLICANT: Lemischka, Ihor R.

TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: ImClone Systems Incorporated

STREET: 180 Varick Street

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10014

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/021,324

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/977,451

FILING DATE: 1992-11-19

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/906,397

FILING DATE: 26-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US PCT/US92/05401

FILING DATE: 26-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: TW 81102961

FILING DATE: 15-APR-1992  
PRIOR APPLICATION DATA: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3453 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 112..3006  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 31..111  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 31..3009  
US-09-021-324-1

Query Match 100.0%; Score 3453; DB 2; Length 3453;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCTGTGCTACCGCGCTCCGGAGGCCATCGGCGGTGGCGCAGCGCAGCGG 60  
Db 1 GCGGCTGTGCTACCGCGCTCCGGAGGCCATCGGCGGTGGCGCAGCGCAGCGG 60  
QY 61 CGGCTGTGCTGTGTTGTTGTCAGTAAATGTTTTCAGACCGTTTACAAACAAG 120  
Db 61 CGGCTGTGCTGTGTTGTTGTCAGTAAATGTTTTCAGACCGTTTACAAACAAG 120  
QY 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATCAGAACAAATGCTCATCAGCGGGAAG 180  
Db 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATCAGAACAAATGCTCATCAGCGGGAAG 180  
QY 181 CCATCATCGTACCGAATGGTCCGAGGATCCCGAGAACCTTCCAGTGTACCCCGAGCGC 240  
Db 181 CCATCATCGTACCGAATGGTCCGAGGATCCCGAGAACCTTCCAGTGTACCCCGAGCGC 240  
QY 241 CAGAGTGAAGGAGCGGTATATGAAGCGGCACCGTGGAGTGGCCGAGTCTGGGTCCATC 300  
Db 241 CAGAGTGAAGGAGCGGTATATGAAGCGGCACCGTGGAGTGGCCGAGTCTGGGTCCATC 300  
QY 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGAGACCTTTTCCCTGCTCTGCTTTTAAAGC 360  
Db 301 ACCCTGCAAGTGCAGCTCGCCACCCAGGAGACCTTTTCCCTGCTCTGCTTTTAAAGC 360  
QY 361 AGCTCCCTGGGCTGCCAGCGCACTTTTATTACAAAACAGAGGAATCGTTTCCATGCGC 420

|||||  
361 AGCTCCCTGGCTGCGCAGCGCACCTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 420  
QY  
421 ATCTTTGAAGCTCACAGAGACCCAGCAGGAGAAATACCTACTCCATATTCAGAGCGAACCC 480  
Db  
421 ATCTTTGAAGCTCACAGAGACCCAGCAGGAGAAATACCTACTCCATATTCAGAGCGAACCC 480  
QY  
481 GCCAACTACACAGTACTCTTACAGTGAATGTAAAGAGATACACAGCTGTATGTCTAAGG 540  
Db  
481 GCCAACTACACAGTACTCTTACAGTGAATGTAAAGAGATACACAGCTGTATGTCTAAGG 540  
QY  
541 AGACCTTACTTTAGGAAGATGAAACAGAGATGCACCTGCTCTGCATCTCCAGGGTGT 600  
Db  
541 AGACCTTACTTTAGGAAGATGAAACAGAGATGCACCTGCTCTGCATCTCCAGGGTGT 600  
QY  
601 CGGAGGCCACTGTGAGTGGGTCTCTGCAGCTCCACAGGAAAGCTGTAAGAGAA 660  
Db  
601 CGGAGGCCACTGTGAGTGGGTCTCTGCAGCTCCACAGGAAAGCTGTAAGAGAA 660  
QY  
661 GGCCCTGCTGTTCAGAAAGAGGAAAGGTACTTTCATGAGTTGTTGCGAAACAGACATC 720  
Db  
661 GGCCCTGCTGTTCAGAAAGAGGAAAGGTACTTTCATGAGTTGTTGCGAAACAGACATC 720  
QY  
721 AGATGCTGTGTAGAAATGCATGTGGCCGCGAATGCACCAAGCTGTTCAACCATAGATCTA 780  
Db  
721 AGATGCTGTGTAGAAATGCATGTGGCCGCGAATGCACCAAGCTGTTCAACCATAGATCTA 780  
QY  
781 AACCAGGCTCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGAAACCTTTGG 840  
Db  
781 AACCAGGCTCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGAAACCTTTGG 840  
QY  
841 ATCAGGTGTAAGGCCATCATGTGAACATGGATTTCGGCTCACCTGGAGCTGGAAGAC 900  
Db  
841 ATCAGGTGTAAGGCCATCATGTGAACATGGATTTCGGCTCACCTGGAGCTGGAAGAC 900  
QY  
901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACTCTCCACAAACAGGACCATG 960  
Db  
901 AAAGCCCTGGAGGAGGCGAGCTACTTTGAGATGAGTACTCTCCACAAACAGGACCATG 960  
QY  
961 ATTCGGATTCTTTGGCCCTTTGTGCTTCCGTGGGAAGAACGACCGGATATTACACC 1020  
Db  
961 ATTCGGATTCTTTGGCCCTTTGTGCTTCCGTGGGAAGAACGACCGGATATTACACC 1020  
QY  
1021 TGCTCTCTCAAAGCACCCAGCCAGTCAGGCTTGGTGAACATCCTAGAAAAAGGTTT 1080  
Db  
1021 TGCTCTCTCAAAGCACCCAGCCAGTCAGGCTTGGTGAACATCCTAGAAAAAGGTTT 1080  
QY  
1081 ATAAACGCTACCGCTCGCAAGAGAGTATGAATTTGACCCGTACGAAAGTTCTGCTTC 1140  
Db  
1081 ATAAACGCTACCGCTCGCAAGAGAGTATGAATTTGACCCGTACGAAAGTTCTGCTTC 1140  
QY  
1141 TCAGTCAGGTTTAAAGCTACCCAGATCCGATCCGATCCGATGATCTTCTCAAGCCCTCA 1200  
Db  
1141 TCAGTCAGGTTTAAAGCTACCCAGATCCGATCCGATCCGATGATCTTCTCAAGCCCTCA 1200  
QY  
1201 TTTCTCTGTAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
Db  
1201 TTTCTCTGTAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
QY  
1261 AAGAACAGCCAGGAGAGTACATATTTATGCAGAAAAATGATGACGCCAGTTTCAACAAA 1320  
Db  
1261 AAGAACAGCCAGGAGAGTACATATTTATGCAGAAAAATGATGACGCCAGTTTCAACAAA 1320  
QY  
1321 ATGTTCCAGCTGTAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCCTCAGCCAGCCAG 1380  
Db  
1321 ATGTTCCAGCTGTAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCCTCAGCCAGCCAG 1380  
QY  
1381 GCGTCCCTGCTCTGATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTGTTCGGAC 1440  
Db  
1381 GCGTCCCTGCTCTGATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTGTTCGGAC 1440  
QY  
1441 AAATCTCCCAATTTGCACGGAGGAAATCCCAAGAGGAGTTTGAATATAAAGGCTAACAGA 1500  
|||||

Db  
1441 AAATCTCCCAATTTGCACGGAGGAAATCCCAAGAGGAGTTTGAATATAAAGGCTAACAGA 1500  
QY  
1501 AAAGTGTTTTGGCCAGTGGGTGTCCAGCAGTACTCTTAATATGATGAGTGGGCGGAAAGGG 1560  
Db  
1501 AAAGTGTTTTGGCCAGTGGGTGTCCAGCAGTACTCTTAATATGATGAGTGGGCGGAAAGGG 1560  
QY  
1561 CTTCTGCTCAAAATGCTGTGCTGTAACAATTTCTATGGGCACGCTTTCGCAAAACCATCTTTT 1620  
Db  
1561 CTTCTGCTCAAAATGCTGTGCTGTAACAATTTCTATGGGCACGCTTTCGCAAAACCATCTTTT 1620  
QY  
1621 AACTCACAGGCCCCCTTCCCTTTTCATCCAAAGACAATCTCTTCTATGCGACCATTTGGG 1680  
Db  
1621 AACTCACAGGCCCCCTTCCCTTTTCATCCAAAGACAATCTCTTCTATGCGACCATTTGGG 1680  
QY  
1681 CTTCTGCTCCCCCTTCAATTTGTTCTCATTTGTTCTGATCTGCCACAATATAAAGCAA 1740  
Db  
1681 CTTCTGCTCCCCCTTCAATTTGTTCTCATTTGTTCTGATCTGCCACAATATAAAGCAA 1740  
QY  
1741 TTTAGTAGAGAGTGCAGCTGCAGATGATCCAGGTGACTGGCCCCCTTGGATAACGAGTAC 1800  
Db  
1741 TTTAGTAGAGAGTGCAGCTGCAGATGATCCAGGTGACTGGCCCCCTTGGATAACGAGTAC 1800  
QY  
1801 TTTACGTTGACTTTCAGGACTATGAATATGACCTTAAAGTGGAGTTCGCGAGAGAAC 1860  
Db  
1801 TTTACGTTGACTTTCAGGACTATGAATATGACCTTAAAGTGGAGTTCGCGAGAGAAC 1860  
QY  
1861 TTTAGAGTTTGGAGAGTCTCTGGGCTGGGGCTTTCGGGAGGTTGATGAAGCCACGGCC 1920  
Db  
1861 TTTAGAGTTTGGAGAGTCTCTGGGCTGGGGCTTTCGGGAGGTTGATGAAGCCACGGCC 1920  
QY  
1921 TATGGCATTAGTAAAAACGGAGTCTCAATTCAGTGGCGGTGAAGTGTCTAAAGAGAAA 1980  
Db  
1921 TATGGCATTAGTAAAAACGGAGTCTCAATTCAGTGGCGGTGAAGTGTCTAAAGAGAAA 1980  
QY  
1981 GCTACAGCTGTGAAAAAGAAAGCTCTCATGTCCGAGCTCAAAATGATGATGACCCACCTGGGA 2040  
Db  
1981 GCTACAGCTGTGAAAAAGAAAGCTCTCATGTCCGAGCTCAAAATGATGATGACCCACCTGGGA 2040  
QY  
2041 CACCATGACAACATCGTGAATCTGCTGGGGCATGCACACTGTCCAGGGCCAGTGTACTTG 2100  
Db  
2041 CACCATGACAACATCGTGAATCTGCTGGGGCATGCACACTGTCCAGGGCCAGTGTACTTG 2100  
QY  
2101 ATTTTGGATATTTGTTGCTATGTTGCTCACTCACTACCTAAAGAGTAAAAAGAGACAAG 2160  
Db  
2101 ATTTTGGATATTTGTTGCTATGTTGCTCACTCACTACCTAAAGAGTAAAAAGAGACAAG 2160  
QY  
2161 TTTACAGGACATGGACAGAGATTTTAAAGAAACATAATTTACAGTCTTTACCCTACTTTC 2220  
Db  
2161 TTTACAGGACATGGACAGAGATTTTAAAGAAACATAATTTACAGTCTTTACCCTACTTTC 2220  
QY  
2221 CAGGCACATTTCAAAATTCAGCATGCTGTTTACAGAGAGTTCAGTTCACCCGCCCTTG 2280  
Db  
2221 CAGGCACATTTCAAAATTCAGCATGCTGTTTACAGAGAGTTCAGTTCACCCGCCCTTG 2280  
QY  
2281 GATCAGCTCTCAGGTTCAATGGGAATTCAAATTCATTCTGAAGATGAGATTGAATGAA 2340  
Db  
2281 GATCAGCTCTCAGGTTCAATGGGAATTCAAATTCATTCTGAAGATGAGATTGAATGAA 2340  
QY  
2341 AACCAGAAGAGCTGGCAGAAAGAGAGAGATTTTGAACGTGCTGACGTTTGAAGAC 2400  
Db  
2341 AACCAGAAGAGCTGGCAGAAAGAGAGAGATTTTGAACGTGCTGACGTTTGAAGAC 2400  
QY  
2401 CTTCTTTGCTTTGCGTACCAAGTGGCCAAAGGATGGAATTTCTTGAGTTCAGTCTGTT 2460  
Db  
2401 CTTCTTTGCTTTGCGTACCAAGTGGCCAAAGGATGGAATTTCTTGAGTTCAGTCTGTT 2460  
QY  
2461 GTCCACAGAGACTGGCAGCCAGGATGTGGTTCACCCACCGGAGGTGGTGAAGATC 2520  
Db  
2461 GTCCACAGAGACTGGCAGCCAGGATGTGGTTCACCCACCGGAGGTGGTGAAGATC 2520  
QY  
2521 TGTGACTTTGGACTGGCCGAGACATCTTCAGGACTCCAGCTACCTGCTCAGGGGCAAC 2580  
Db  
2521 TGTGACTTTGGACTGGCCGAGACATCTTCAGGACTCCAGCTACCTGCTCAGGGGCAAC 2580



Db 481 GCCAACTACACAGTACTGTTTACACAGTGAATGTAAAGAGATACACAGCTGTATGTCTAAGG 540  
Qy 541 AGACCTTACTTTAGGAGATGGAACACAGATGCACTGCTCTCCATCTCCGAGGGTGT 600  
Db 541 AGACCTTACTTTAGGAGATGGAACACAGATGCACTGCTCTCCATCTCCGAGGGTGT 600  
Qy 601 CCGAGGCCACTGTGGAGTGGGTGCTCTGCAAGTCCACAGGAAAGCTGTAAAGAAAGAA 660  
Db 601 CCGAGGCCACTGTGGAGTGGGTGCTCTGCAAGTCCACAGGAAAGCTGTAAAGAAAGAA 660  
Qy 661 GGCCCTGCTGTTGTCAGAAAGAGGAAAGTACTTCAATGAGTTGTCGGAACAGACATC 720  
Db 661 GGCCCTGCTGTTGTCAGAAAGAGGAAAGTACTTCAATGAGTTGTCGGAACAGACATC 720  
Qy 721 AGATGCTGCTGAATGACACTGGCCGCGAATGACCAAGCTGTTACCATAGATCTA 780  
Db 721 AGATGCTGCTGAATGACACTGGCCGCGAATGACCAAGCTGTTACCATAGATCTA 780  
Qy 781 AACACAGGCTCCTCAGACACACTGCCCCAGTATTCTGAAAGTGGGGAAACCTTTGTGG 840  
Db 781 AACACAGGCTCCTCAGACACACTGCCCCAGTATTCTGAAAGTGGGGAAACCTTTGTGG 840  
Qy 841 ATCAGGTGTAAAGCCATCCATGTAACCATGGATTTCGGGCTCACCTGGGAGCTGGAAGAC 900  
Db 841 ATCAGGTGTAAAGCCATCCATGTAACCATGGATTTCGGGCTCACCTGGGAGCTGGAAGAC 900  
Qy 901 AAAGCCCTGAGGAGGCGACTACTTTGAGATGAGTACTTCCACAAACAGGACCATG 960  
Db 901 AAAGCCCTGAGGAGGCGACTACTTTGAGATGAGTACTTCCACAAACAGGACCATG 960  
Qy 961 ATTGCGATTCCTTTGGCTTTGTTGCTTCCGTGGGAAGGACGACCGGATATTACACC 1020  
Db 961 ATTGCGATTCCTTTGGCTTTGTTGCTTCCGTGGGAAGGACGACCGGATATTACACC 1020  
Qy 1021 TGCTCTTCTCAAAGCACCAGCCAGTCAAGCTGTGTGACCATCTAGACCAAGGTTT 1080  
Db 1021 TGCTCTTCTCAAAGCACCAGCCAGTCAAGCTGTGTGACCATCTAGACCAAGGTTT 1080  
Qy 1081 ATAAAGCTACAGCTCGAAGAGAGTATGAATTCACCGGTACGAAAGTTCGTGCTC 1140  
Db 1081 ATAAAGCTACAGCTCGAAGAGAGTATGAATTCACCGGTACGAAAGTTCGTGCTC 1140  
Qy 1141 TCAGTCAAGTTTAAAGCGTACCCAGAAATCCGATGACGCTGGATCTCTCAAGCCCTCA 1200  
Db 1141 TCAGTCAAGTTTAAAGCGTACCCAGAAATCCGATGACGCTGGATCTCTCAAGCCCTCA 1200  
Qy 1201 TTTCTCTTGAACAGAGGCGCTGGAGATGGGTACAGCATATCTAAATTTTGGATCAT 1260  
Db 1201 TTTCTCTTGAACAGAGGCGCTGGAGATGGGTACAGCATATCTAAATTTTGGATCAT 1260  
Qy 1261 AAGAACAGCCAGGAGTACATATCTATGCAAGAAATGATGACGCGCCAGTTTCAACAAA 1320  
Db 1261 AAGAACAGCCAGGAGTACATATCTATGCAAGAAATGATGACGCGCCAGTTTCAACAAA 1320  
Qy 1321 ATGTTCAAGCTGATATAAGAAACAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG 1380  
Db 1321 ATGTTCAAGCTGATATAAGAAACAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG 1380  
Qy 1381 GCGTCCTGTTCTCTGATGGCTACCCGCTACCCCTTCTGGACCTGGAAGAGTGTTCGGAC 1440  
Db 1381 GCGTCCTGTTCTCTGATGGCTACCCGCTACCCCTTCTGGACCTGGAAGAGTGTTCGGAC 1440  
Qy 1441 AAATCTCCCAATTCACGAGGAAATCCAGAGAGGTTTGGAAATAAAGGCTTAACAGA 1500  
Db 1441 AAATCTCCCAATTCACGAGGAAATCCAGAGAGGTTTGGAAATAAAGGCTTAACAGA 1500  
Qy 1501 AAAGTGTGTCAGTGGGTTCAGCAGTACTTAAATATGAGTGGCCGCGGAAAGGG 1560  
Db 1501 AAAGTGTGTCAGTGGGTTCAGCAGTACTTAAATATGAGTGGCCGCGGAAAGGG 1560  
Qy 1561 CTTCTGCTCAATCTGTGCTTCAATTTCTATGGCAGCTTTCGCAACCATCTTTTAA 1620  
Db 1561 CTTCTGCTCAATCTGTGCTTCAATTTCTATGGCAGCTTTCGCAACCATCTTTTAA 1620

Qy 1621 AACTCACAGGCCCTTCCCTTTTCATCAAGACAAACATCTCTTCTATGCGACCATTTGG 1680  
Db 1621 AACTCACAGGCCCTTCCCTTTTCATCAAGACAAACATCTCTTCTATGCGACCATTTGG 1680  
Qy 1681 CTCTGTCTCCCTTCATTTGTTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740  
Db 1681 CTCTGTCTCCCTTCATTTGTTTCTCATTTGTTGATCTGCCACAAATACAAAAGCAA 1740  
Qy 1741 TTTAGGTACGAGAGTCACTGTCAGATGATCAGGTGACTGGCCCTCGATTAACGAGTAC 1800  
Db 1741 TTTAGGTACGAGAGTCACTGTCAGATGATCAGGTGACTGGCCCTCGATTAACGAGTAC 1800  
Qy 1801 TTTACGTTGACTTTCAGGAGTATGAATATACCTTTAAGTGGAGTTCCTCCGAGAGAAC 1860  
Db 1801 TTTACGTTGACTTTCAGGAGTATGAATATACCTTTAAGTGGAGTTCCTCCGAGAGAAC 1860  
Qy 1861 TTAGAGTTTGGGAAGTCTTGGGCTTCTGGGCTTTCGGGAGGTTGATGAACGCCAGGCC 1920  
Db 1861 TTAGAGTTTGGGAAGTCTTGGGCTTCTGGGCTTTCGGGAGGTTGATGAACGCCAGGCC 1920  
Qy 1921 TATGGCATTAAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA 1980  
Db 1921 TATGGCATTAAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTTAAAGAGAAA 1980  
Qy 1981 GCTCACAGCTGTGAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040  
Db 1981 GCTCACAGCTGTGAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040  
Qy 2041 CACCATGACACATCTGTAATCTGCTGGGCGATGACACACTGTACAGGCCAGTGTACTTG 2100  
Db 2041 CACCATGACACATCTGTAATCTGCTGGGCGATGACACACTGTACAGGCCAGTGTACTTG 2100  
Qy 2101 ATTTTGAATATTTGCTGATGTTGCTGATGTTGCTCAACTACCTAAGAGTAAAGAGAGAG 2160  
Db 2101 ATTTTGAATATTTGCTGATGTTGCTGATGTTGCTCAACTACCTAAGAGTAAAGAGAGAG 2160  
Qy 2161 TTTTACAGGACATGACAGAGATTTTAAAGAAACATAATTTTCACTTCACTTACTTTC 2220  
Db 2161 TTTTACAGGACATGACAGAGATTTTAAAGAAACATAATTTTCACTTCACTTACTTTC 2220  
Qy 2221 CAGCAGATTCACAAATTCAGCATGCTGTTTACAGAGAGTTCAGTTTACACCCGCTTTC 2280  
Db 2221 CAGCAGATTCACAAATTCAGCATGCTGTTTACAGAGAGTTCAGTTTACACCCGCTTTC 2280  
Qy 2281 GATCAGCTCTCAGGGTTCAATTTGGAATTCATTTCTGAAGATGAGATTGAATGAA 2340  
Db 2281 GATCAGCTCTCAGGGTTCAATTTGGAATTCATTTCTGAAGATGAGATTGAATGAA 2340  
Qy 2341 AACCAGAGAGGCTGGCAGAGAGAGAGAGAGATTTTGAACGCTGCTGACGTTTGAAGAC 2400  
Db 2341 AACCAGAGAGGCTGGCAGAGAGAGAGAGAGATTTTGAACGCTGCTGACGTTTGAAGAC 2400  
Qy 2401 CTCCTTTGCTTTGCTTACCAAGTGGCAAGAGCATGGAATTCCTGGAGTTCAAGTCTGT 2460  
Db 2401 CTCCTTTGCTTTGCTTACCAAGTGGCAAGAGCATGGAATTCCTGGAGTTCAAGTCTGT 2460  
Qy 2461 GTCCACAGAGACCTGGCAGCAGGAATGTTGTTGTCACCCAGGGAAGTGGTGAAGATC 2520  
Db 2461 GTCCACAGAGACCTGGCAGCAGGAATGTTGTTGTCACCCAGGGAAGTGGTGAAGATC 2520  
Qy 2521 TGTGACTTTGACTGGCCCGGAGACATCTGAGCGACTCCAGCTAGCTGCTCAGGCGCAAC 2580  
Db 2521 TGTGACTTTGACTGGCCCGGAGACATCTGAGCGACTCCAGCTAGCTGCTCAGGCGCAAC 2580  
Qy 2581 GCAGGCTGGCGGTGAAGTGGATGGATGGCCACCCGAGAGCTTATTGAGGGATCTACAAATC 2640  
Db 2581 GCAGGCTGGCGGTGAAGTGGATGGATGGCCACCCGAGAGCTTATTGAGGGATCTACAAATC 2640  
Qy 2641 AAGAGTACGCTGCTGCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700  
Db 2641 AAGAGTACGCTGCTGCTTACGGCATCTTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700



Qy	661	GGCCCTGCTGTTGTGCAGAAAGAGGAAAAGGTACTTTCATGAGTTGTTCTCGGAACACACATC	720	Qy	1741	TTTAGGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCCCTCGATACACGATAC	1800
Db	661		720	Db	1741		1800
Qy	721	AGATGCTGTGTAGAAATGCATCGGCGCGGAAATGCACCAAGCTGTTCACCATAGATCTA	780	Qy	1801	TTTACGTTGACTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCGCCAGAGAGAAC	1860
Db	721		780	Db	1801		1860
Qy	781	AACCAAGCTCCTCAGAGCAGACTGCCCCAGCTATTCCTGAAAGTGGGGGAAACCTTTGCG	840	Qy	1861	TTAGAGTTTGGGAAGGCTCTGGGCTTTCGGAGGGTGAATGACGCCACCGGCC	1920
Db	781		840	Db	1861		1920
Qy	841	ATCAGGTTGAAGGCATCCATGTGACCATGGATTTCGGGCTCACCTGGAGCTGGAAGAC	900	Qy	1921	TATGCATTAGTAAACCGGAGCTCAATTCAGGTGGCGGTGAAGTGTAAAGAGAAA	1980
Db	841		900	Db	1921		1980
Qy	901	AAAGCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTACTCCAAACAGGACCATTG	960	Qy	1981	GCTGACAGCTGTGAAAAGAAAGCTCTCATGTGCGGAGCTCAAAATGATGACCCACCTGGGA	2040
Db	901		960	Db	1981		2040
Qy	961	ATTCCGATTCTTTCGGCTTTGTCTTCCTGGGAGGAGGACGACCGGATATTACACC	1020	Qy	2041	CACCATGACACATCGTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTGTACTTG	2100
Db	961		1020	Db	2041		2100
Qy	1021	TGCTCTTCTCAAAGCACCCAGCAGTCAGCGTTGGTGACCATCTAGAAAAAGGGTTT	1080	Qy	2101	ATTTTGAATATTGTTGCTATGTGACCTCCCTCACTACCTTAAGAAAGTAAAGAGAGAAG	2160
Db	1021		1080	Db	2101		2160
Qy	1081	ATAAACGCTACAGCTCGCAAGAGAGTATGAAATTTGACCGGTACGAAAAGTCTTGCTTC	1140	Qy	2161	TTTACAGAGACATGGACAGAGATTTTAAAGAACTAAATTCAGTTCTTACCCTACTTTC	2220
Db	1081		1140	Db	2161		2220
Qy	1141	TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATCGATGCGATGCTTCTCAAGCCCTCA	1200	Qy	2221	CAGGCACATTCAAATTCACGATGCTGTTTACAGGAGAGTTCAGTTACACCCGCCCTTG	2280
Db	1141		1200	Db	2221		2280
Qy	1201	TTTCTCTGTGAACAGAGGCTCGAGGATGGGTACACATATCTAAATTTTGGCATCAT	1260	Qy	2281	GATCAGCTCTCAGGTTCAATGGGAATTCATTCATTCTGAAGATGAGATTGAATGAA	2340
Db	1201		1260	Db	2281		2340
Qy	1261	AAGAACAGCAGAGAGTACATATTCATGCAGAAAATGATGAGCCGACGTTTACACAAA	1320	Qy	2341	AACCAAGAGAGCTGGCAGAAAGAGAGGAGGAGGATTTGAACGTGCTGACGTTTGAAGAC	2400
Db	1261		1320	Db	2341		2400
Qy	1321	ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGAAAATGCTCAGCCAGCCAG	1380	Qy	2401	CTCCTTTGCTTTGCTTACCAAGTGGCCAAAGGATGGAATTCCTGAGTTCAAGTCGTGT	2460
Db	1321		1380	Db	2401		2460
Qy	1381	CGTCTCTTCTCTGATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTGTTCGGAC	1440	Qy	2461	GTCCACAGAGACTGGCAGCCAGGATGTGTTGGTCAACCCACGGGAGGTGGTGAAGATC	2520
Db	1381		1440	Db	2461		2520
Qy	1441	AAATCTCCCAATTCACAGAGGAAATCCAGAAAGGTTTGAATATAAGAGGCTAACAGA	1500	Qy	2521	TGTGACTTTGGACTGGCCCGAGACATCCTGAGCGACTCCAGCTACGTTCTGAGGGGCAAC	2580
Db	1441		1500	Db	2521		2580
Qy	1501	AAAGTGTTCGCGAGTGGGTGAGCAGTACTCTAAATATGAGTGAGGCGCGGAAAGGG	1560	Qy	2581	GCACGGCTGCCGGTGAAGTGGATGGCACCAGAGCTTATTTGAAGGGATCTACACAATC	2640
Db	1501		1560	Db	2581		2640
Qy	1561	CTTCTGGTCAATGCTGTGCTACAAATCTATGGCAGCTCTTGGAAACCATCTTTTAA	1620	Qy	2641	AAGAGTGAGCTCTGGTCTTACGGCATCTTCTCTGGGAGATATTTTACCTGGGTGTGAAC	2700
Db	1561		1620	Db	2641		2700
Qy	1621	AACTCACCAGGCCCTTCCCTTTATCCAAAGACACATCTCCTCTATGCGACCATTTGGG	1680	Qy	2701	CCTTACCCTGGCATCTCTGACCGCTAACTTCTATAAACTGATTACAGAGTGGATTTAAA	2760
Db	1621		1680	Db	2701		2760
Qy	1681	CTGTCTCCCTTCAATGTTGTTCTCATTTGTTGATCTGCGCAAAATACAAAAAGCAA	1740	Qy	2761	ATGAGCAGCCATCTATGCCACAGAAAGGATATACTTTGTAATGCAATCCTCTGGGCT	2820
Db	1681		1740	Db	2761		2820
Qy	1741	CTGTCTCCCTTCAATGTTGTTCTCATTTGTTGATCTGCCACAAATACAAAAAGCAA	1800	Qy	2821	TTTGACTCAAGGAAGGGGCCATCCTTCCCAACCTGACTTCATTTTAGGATGTCAGCTG	2880
Db	1741		1800	Db	2821		2880





Db 661 GGCCCTGCTGTTGTGTCAGAAAGGAGGAAAGGTAAGTCTTTCATGAGTTGTTCCGGAAACAGACATC 720  
Qy 721 AGATGCTGTGCTAGAAATGCACTGGCGCGCAANTGCCAACAGCTGTTCCACCATAGATCTA 780  
Db 721 ATATGCTGTGCTAGAAATGCACTGGCGCGCAANTGCCAACAGCTGTTCCACCATAGATCTA 780  
Qy 781 AACCAGGCTCCTCAGAGCACACTGCCCCAGTTATTCTCTGAAAGTGGGGAACCCCTTGCG 840  
Db 781 AACCAGGCTCCTCAGAGCACACTGCCCCAGTTATTCTCTGAAAGTGGGGAACCCCTTGCG 840  
Qy 841 ATCAGGTGTAGGCCATCCATGTGAACCATGGATTTCGGGCTCACCTGGGAGCTGGAAGAC 900  
Db 841 ATCAGGTGTAGGCCATCCATGTGAACCATGGATTTCGGGCTCACCTGGGAGCTGGAAGAC 900  
Qy 901 AAACCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTACTCCACAAACAGACCATG 960  
Db 901 AAACCCCTGGAGGAGGAGCTACTTTGAGATGAGTACTACTCCACAAACAGACCATG 960  
Qy 961 ATTCCGATTCTCTTGGCCCTTTGTCTTCCGTGGGAAGGAACGACACCGGATATTACACC 1020  
Db 961 ATTCCGATTCTCTTGGCCCTTTGTCTTCCGTGGGAAGGAACGACACCGGATATTACACC 1020  
Qy 1021 TGCTCTTCCCTCAAAAGCACCCAGCAGTCAAGCTGGTGGTACCATCTTAGAAAAAGGGTTT 1080  
Db 1021 TGCTCTTCCCTCAAAAGCACCCAGCAGTCAAGCTGGTGGTACCATCTTAGAAAAAGGGTTT 1080  
Qy 1081 ATAAACGCTACCGCTCGCAGAGAGATATCAAAATTGACCCGTACGAAAGCTTCTGCTTC 1140  
Db 1081 ATAAACGCTACCGCTCGCAGAGAGATATCAAAATTGACCCGTACGAAAGCTTCTGCTTC 1140  
Qy 1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATGCGATGCTCTCTCAAGCCCTCA 1200  
Db 1141 TCAGTCAGGTTTAAAGCGTACCCAGCAATCCGATGCGATGCTCTCTCAAGCCCTCA 1200  
Qy 1201 TTTCCCTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTGGGATCAT 1260  
Db 1201 TTTCCCTGTGAACAGAGAGGCTGGAGATGGGTACAGCATATCTAAATTTGGGATCAT 1260  
Qy 1261 AAGAACAGGCGAGAGAGTACATATCTATCAGAAATGATGAGCCAGCTTACCAAAA 1320  
Db 1261 AAGAACAGGCGAGAGAGTACATATCTATCAGAAATGATGAGCCAGCTTACCAAAA 1320  
Qy 1321 ATGTTACCGTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCCCTCAGCCAGCCAG 1380  
Db 1321 ATGTTACCGTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCCCTCAGCCAGCCAG 1380  
Qy 1381 CGGTCCTGTTCTCTGATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTGTTCGGAC 1440  
Db 1381 CGGTCCTGTTCTCTGATGGCTACCCGCTACCCCTTGGACCTGGAAGAGTGTTCGGAC 1440  
Qy 1441 AAATCTCCAAATTCACGGAGGAATCCAGAAAGAGTTTGGATTAAGAGCTTAACAGA 1500  
Db 1441 AAATCTCCAAATTCACGGAGGAATCCAGAAAGAGTTTGGATTAAGAGCTTAACAGA 1500  
Qy 1501 AAAGTGTGGCCAGTGGGTCTGAGCAGTACTCTAAATATGATGAGGCGCGGAAAGGG 1560  
Db 1501 AAAGTGTGGCCAGTGGGTCTGAGCAGTACTCTAAATATGATGAGGCGCGGAAAGGG 1560  
Qy 1561 CTTCGTGTCAAATGCTGTGCTACAAATCTATGGCAGCTCTTCGAAACCATCTTTTA 1620  
Db 1561 CTTCGTGTCAAATGCTGTGCTACAAATCTATGGCAGCTCTTCGAAACCATCTTTTA 1620  
Qy 1621 AACTCAGCGCCCTTCCCTTTTCATCCAGACAAACATCTCTTCTATGCGACCATTTGG 1680  
Db 1621 AACTCAGCGCCCTTCCCTTTTCATCCAGACAAACATCTCTTCTATGCGACCATTTGG 1680  
Qy 1681 CTCTGCTCCCTTCATTTGTTGTTCTCATTTGTTGATCTGCCCAATAACAAAGCAA 1740  
Db 1681 CTCTGCTCCCTTCATTTGTTGTTCTCATTTGTTGATCTGCCCAATAACAAAGCAA 1740  
Qy 1741 TTTAGGTACGAGGTACAGTGCAGATGATCCAGGTGACTGGCCCTCGGATTAACAGTAC 1800  
Db 1741 TTTAGGTACGAGGTACAGTGCAGATGATCCAGGTGACTGGCCCTCGGATTAACAGTAC 1800

Qy 1801 TTCTACGTTGACTTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAGAAC 1860  
Db 1801 TTCTACGTTGACTTTCAGGAGCTATGAATATGACCTTAAGTGGAGTTCCCGAGAGAGAAC 1860  
Qy 1861 TTAGAGTTTGGGAAGGTCTTGGGCTTGGGAGGTGATGAACGCCACGGCC 1920  
Db 1861 TTAGAGTTTGGGAAGGTCTTGGGCTTGGGAGGTGATGAACGCCACGGCC 1920  
Qy 1921 TATGCAATAGTAAACGGGAGTCTCAATTCAGTGGCGGTGAAGATGTAAAGAGAAA 1980  
Db 1921 TATGCAATAGTAAACGGGAGTCTCAATTCAGTGGCGGTGAAGATGTAAAGAGAAA 1980  
Qy 1981 GCTGACAGCTGTGAAAAAGAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA 2040  
Db 1981 GCTGACAGCTGTGAAAAAGAGCTCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGA 2040  
Qy 2041 CACCATGACAATCTGCTGCTGGGCGATGCACACTGTACAGGCGAGTGTACTTG 2100  
Db 2041 CACCATGACAATCTGCTGCTGGGCGATGCACACTGTACAGGCGAGTGTACTTG 2100  
Qy 2101 ATTTTGAATATTTGTTGCTATGTTGACCTCTCAACTACCTAAGAGTAAAGAGAGAAG 2160  
Db 2101 ATTTTGAATATTTGTTGCTATGTTGACCTCTCAACTACCTAAGAGTAAAGAGAGAAG 2160  
Qy 2161 TTTACAGAGCATGCACAGAGATTTTAAAGAACATAATTTTTCAGTTCTTACCCTACTTTC 2220  
Db 2161 TTTACAGAGCATGCACAGAGATTTTAAAGAACATAATTTTTCAGTTCTTACCCTACTTTC 2220  
Qy 2221 CAGGACATTTCAAATTCAGCATGCTGTTTCCAGGAGAGTTCAGTTACACCCGCCCTTG 2280  
Db 2221 CAGGACATTTCAAATTCAGCATGCTGTTTCCAGGAGAGTTCAGTTACACCCGCCCTTG 2280  
Qy 2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTCATCTGAAAGTGAAGTGAATGAA 2340  
Db 2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTCATCTGAAAGTGAAGTGAATGAA 2340  
Qy 2341 AACCAAGAGAGCTGGCAGAAAGAGAGAGATTTTGAACGCTGCTGACGTTTGAAGAC 2400  
Db 2341 AACCAAGAGAGCTGGCAGAAAGAGAGAGATTTTGAACGCTGCTGACGTTTGAAGAC 2400  
Qy 2401 CTCTTTTCTGCTTACCAAGTGGCCAAAGGAGTTCCTGGAGTTCAAGTCTGT 2460  
Db 2401 CTCTTTTCTGCTTACCAAGTGGCCAAAGGAGTTCCTGGAGTTCAAGTCTGT 2460  
Qy 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGTCACCCAGGGAAGTGTGAAGATC 2520  
Db 2461 GTCCACAGAGACCTGGCAGCCAGGAATGTGTTGTCACCCAGGGAAGTGTGAAGATC 2520  
Qy 2521 TGTGACTTTGAGTGGCGCGGAGACATCCGAGGACTCCAGCTACGTCGAGGCGAAC 2580  
Db 2521 TGTGACTTTGAGTGGCGCGGAGACATCCGAGGACTCCAGCTACGTCGAGGCGAAC 2580  
Qy 2581 GCACGGCTGCCGGTGAAGTGGATGGCACCAGAGAGCTTATTTGAAGGATCTACACAATC 2640  
Db 2581 GCACGGCTGCCGGTGAAGTGGATGGCACCAGAGAGCTTATTTGAAGGATCTACACAATC 2640  
Qy 2641 AAGAGTACGCTGTGTTCTCTACGGCATCTCTCTGAGAGATATTTTCACTGGGTGTGAAC 2700  
Db 2641 AAGAGTACGCTGTGTTCTCTACGGCATCTCTCTGAGAGATATTTTCACTGGGTGTGAAC 2700  
Qy 2701 CCTTACCCTGGCATTCCTGTGACGCTTCTATATAACTGATTCAGAGTGGATTTAA 2760  
Db 2701 CCTTACCCTGGCATTCCTGTGACGCTTCTATATAACTGATTCAGAGTGGATTTAA 2760  
Qy 2761 ATGGAGCAGCATTCATGCGCACAGAGGATATCTTTGTAATGCAATCTGCTGGCT 2820  
Db 2761 ATGGAGCAGCATTCATGCGCACAGAGGATATCTTTGTAATGCAATCTGCTGGCT 2820  
Qy 2821 TTTGACTCAAGAGAGCGGCCATCTTCCCAACCTGACTTCAATTTTAGGATGTACGCTG 2880  
Db 2821 TTTGACTCAAGAGAGCGGCCATCTTCCCAACCTGACTTCAATTTTAGGATGTACGCTG 2880



Qy	181	CCATCATCGTACCGAATGFGCGCAGGATCCCCAGAAGAACTCCAGTGTATCCCCGAGGCGC	240
Db	232		
		CCATCATCGTACCGAATGFGCGCAGGATCCCCAGAAGAACTCCAGTGTATCCCCGAGGCGC	291
Qy	241	CAGAGTGAAGGGACGGTATATGAAGCGGCCACCGTGGAGTGGCCAGTCTGGGTCCATC	300
Db	292		
		CAGAGTGAAGGGACGGTATATGAAGCGGCCACCGTGGAGTGGCCAGTCTGGGTCCATC	351
Qy	301	ACCTTGAAGTGCAGCTCGCCACCCAGGCGACTTTCTCTCCCTCTGGTCTTTTAAAGCAC	360
Db	352		
		ACCTTGAAGTGCAGCTCGCCACCCAGGCGACTTTCTCTCCCTCTGGTCTTTAAGCAC	411
Qy	361	AGCTCCCTGGGCTGCCAGCGCGACTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC	420
Db	412		
		AGCTCCCTGGGCTGCCAGCGCGACTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC	471
Qy	421	ATCTTGAACGTGACAGACCCAGGAGAGAACTACTCTCCATATTAGAGCGAAGCC	480
Db	472		
		ATCTTGAACGTGACAGACCCAGGAGAGAACTACTCTCCATATTAGAGCGAAGCC	531
Qy	481	GCCAACTACACAGTACTGTTACAGTGAATGTAAAGATACACAGCTGTATGTCTAAAG	540
Db	532		
		GCCAACTACACAGTACTGTTACAGTGAATGTAAAGATACACAGCTGTATGTCTAAAG	591
Qy	541	AGACCTTACTTTAGGAAGATGGAACACGAGATGCACCTGCTGCAATCTCCGAGGGTGT	600
Db	592		
		AGACCTTACTTTAGGAAGATGGAACACGAGATGCACCTGCTGCAATCTCCGAGGGTGT	651
Qy	601	CCGAGGCCACTGTGGAGTGGTGCTCGCAGCTCCCAGAGGAAGCTGTAAGAAGAA	660
Db	652		
		CCGAGGCCACTGTGGAGTGGTGCTCGCAGCTCCCAGAGGAAGCTGTAAGAAGAA	711
Qy	661	GGCCCTGCTGTTGTCAGAAAGAGGAAAAGGTACTTCAATGAGTGTTCGGAACAGACATC	720
Db	712		
		GGCCCTGCTGTTGTCAGAAAGAGGAAAAGGTACTTCAATGAGTGTTCGGAACAGACATC	771
Qy	721	AGATGCTGTCTAGAAATGCACGTGGGCCGGAAATGCACCAAGCTGTTTCAACATAGATCTA	780
Db	772		
		AGATGCTGTCTAGAAATGCACGTGGGCCGGAAATGCACCAAGCTGTTTCAACATAGATCTA	831
Qy	781	AACCAAGCTCTCAGAGCACACTGCCCACTTATCTGAAAGTGGGGGAACCTTGTGG	840
Db	832		
		AACCAAGCTCTCAGAGCACACTGCCCACTTATCTGAAAGTGGGGGAACCTTGTGG	891
Qy	841	ATCAGGTGTAAGGCCATCCATGTGAACCATGGAATTCGGGCTCACCTGGAGCTGGAAGAC	900
Db	892		
		ATCAGGTGTAAGGCCATCCATGTGAACCATGGAATTCGGGCTCACCTGGAGCTGGAAGAC	951
Qy	901	AAAGCCCTGAGGAGGCGACGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG	960
Db	952		
		AAAGCCCTGAGGAGGCGACGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG	1011
Qy	961	ATTCCGATTTCTTGGCCTTTGTCTCTCCGTGGGAAGGAACACACCCGGATATTACACC	1020
Db	1012		
		ATTCCGATTTCTTGGCCTTTGTCTCTCCGTGGGAAGGAACACACCCGGATATTACACC	1071
Qy	1021	TGCTCTTCTTCAAAGCACCCCGACGTCAGCGTTGGTGACCATCTCTAGAAAAAGGGTTT	1080
Db	1072		
		TGCTCTTCTTCAAAGCACCCCGACGTCAGCGTTGGTGACCATCTCTAGAAAAAGGGTTT	1131
Qy	1081	ATAACGCTACCGCTCGCAAGAGAGTATGAAATTGACCCGTACGAAAAAGTTCTGCTTC	1140
Db	1132		
		ATAACGCTACCGCTCGCAAGAGAGTATGAAATTGACCCGTACGAAAAAGTTCTGCTTC	1191
Qy	1141	TCAGTCAAGTTTAAAGCGTACCCACGAAATCCGATGCACGTGGATCTTCTCTCAAGCCTCA	1200
Db	1192		
		TCAGTCAAGTTTAAAGCGTACCCACGAAATCCGATGCACGTGGATCTTCTCTCAAGCCTCA	1251
Qy	1201	TTTCTCTTGAACAGAGAGGCGCTGGAGGATGGGTACAGCATATCTAAATTTTCCGATCAT	1260
Db	1252		
		TTTCTCTTGAACAGAGAGGCGCTGGAGGATGGGTACAGCATATCTAAATTTTCCGATCAT	1311
Qy	1261	AAGAAACAGCCAGGAGAGTACATATTCTATGCAGAAAATGATACGCGCCAGTTTCACCAA	1320

Db	1312	 AAGAACAGCCAGGAGAGTACATATTCTATTCAGAAATGATACGCCCGCCAGTTTCACCAA	1371
Qy	1321	ATGTTTCACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCAGCCAGCCAG	1380
Db	1372	ATGTTTCACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCAGCCAGCCAG	1431
Qy	1381	CGGTCCTGTTCCCTCTGATGCTACCCGCTACCCCTCTTTGGACCTTGGAAAGAGTGTTCGGAC	1440
Db	1432	CGGTCCTGTTCCCTCTGATGCTACCCGCTACCCCTCTTTGGACCTTGGAAAGAGTGTTCGGAC	1491
Qy	1441	AAATCTCCCAATTCACGGAGGAATCCACAGAGAGTGTGGATATAAAGGCGCTAACAGA	1500
Db	1492	AAATCTCCCAATTCACGGAGGAATCCACAGAGAGTGTGGATATAAAGGCGCTAACAGA	1551
Qy	1501	AAAGTGTTTGGCCAGTGGGTGTCCAGCAGTACTCTAAATATGAGTAGGCGCGGGAAGAGG	1560
Db	1552	AAAGTGTTTGGCCAGTGGGTGTCCAGCAGTACTCTAAATATGAGTAGGCGCGGGAAGAGG	1611
Qy	1561	CTTCTGTGCTAAATGCTGTGGTACAAATCTATGGGCACGCTTTCGGCAACCATCTTTTTA	1620
Db	1612	CTTCTGTGCTAAATGCTGTGGTACAAATCTATGGGCACGCTTTCGGCAACCATCTTTTTA	1671
Qy	1621	AACTCACCGGCCCTTCCCTTTCATCCACAGACAACTCTCCTTCTATGCGACCATTTGGG	1680
Db	1672	AACTCACCGGCCCTTCCCTTTCATCCACAGACAACTCTCCTTCTATGCGACCATTTGGG	1731
Qy	1681	CTCTGTCTCCCTTTCATTTGTTCTCATTTGTTGATCTGCCACAATACAAAAAGCAA	1740
Db	1732	CTCTGTCTCCCTTTCATTTGTTCTCATTTGTTGATCTGCCACAATACAAAAAGCAA	1791
Qy	1741	TTTAGGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGSCCCCCCTGGATAACGAGTAC	1800
Db	1792	TTTAGGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGSCCCCCCTGGATAACGAGTAC	1851
Qy	1801	TTTCACGTTGACTTCAGGGACTATGAATATGACCTTAAGTGGGAGTTCCCGAGAGAGAAC	1860
Db	1852	TTTCACGTTGACTTCAGGGACTATGAATATGACCTTAAGTGGGAGTTCCCGAGAGAGAAC	1911
Qy	1861	TTACAGTTTGGGAAGGTCCCTGGGCTCTGGGCTTTCCGGAGGTTGTAACGCCACCGCC	1920
Db	1912	TTACAGTTTGGGAAGGTCCCTGGGCTCTGGGCTTTCCGGAGGTTGTAACGCCACCGCC	1971
Qy	1921	TATGGCATTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAAAGAGAA	1980
Db	1972	TATGGCATTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAAAGAGAA	2031
Qy	1981	GCTGACAGCTGTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA	2040
Db	2032	GCTGACAGCTGTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA	2091
Qy	2041	CACCATGACACATCGTGAATCTGCTGGGGCATGCACATGTCAGGGCCAGTGTACTTGG	2100
Db	2092	CACCATGACACATCGTGAATCTGCTGGGGCATGCACATGTCAGGGCCAGTGTACTTGG	2151
Qy	2101	ATTTTGTGAATATTGCTATGTTGACCTCTCAACTACCTTAAGAAAGTAAAGAGAGAAG	2160
Db	2152	ATTTTGTGAATATTGCTATGTTGACCTCTCAACTACCTTAAGAAAGTAAAGAGAGAAG	2211
Qy	2161	TTTCACAGGACATGGACAGAGATTTTAAAGGAACATAAATTTACCTTCTTACCCTACTTC	2220
Db	2212	TTTCACAGGACATGGACAGAGATTTTAAAGGAACATAAATTTACCTTCTTACCCTACTTC	2271
Qy	2221	CAGGCACATTCAAATTCACGATGCCTGGTTTCAGGAGAGTTCAGTTACACCCGCCCTTG	2280
Db	2272	CAGGCACATTCAAATTCACGATGCCTGGTTTCAGGAGAGTTCAGTTACACCCGCCCTTG	2331
Qy	2281	GATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATCTGAAGATGAGATTGAATATGAA	2340
Db	2332	GATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATCTGAAGATGAGATTGAATATGAA	2391
Qy	2341	AACCAAGAGAGGCTGGCAGAAAGAGGAGGAAGATTGAACGTGCTGACGCTTTGAAGC	2400

2392	Db	AACCAAGAAGCGCTGGCAGAGAAGAAGAGGAAGATTGAACTGCTGAGCGTTTGAGAC	2415
2401	Qy	CTCCTTTGCTTTGCGTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCGTGT	2460
2452	Db	CTCCTTTGCTTTGCGTACCAAGTGGCCAAAGGCATGGAATTCCTGGAGTTCAAGTCGTGT	2511
2461	Qy	GTCCACAGACCTGGCAGCCAGGATGTGTTGGTCACCCACGGGAAGTGGTGAAGATC	2520
2512	Db	GTCCACAGACCTGGCAGCCAGGATGTGTTGGTCACCCACGGGAAGTGGTGAAGATC	2571
2521	Qy	TGTGACTTTGGACTGGCCCGGAGACATCCTCAGCGACTCCAGCTACGTCGACGGGCAAC	2580
2572	Db	TGTGACTTTGGACTGGCCCGGAGACATCCTCAGCGACTCCAGCTACGTCGACGGGCAAC	2631
2581	Qy	GCACGGCTGCCGTGAAGTGGATGGCACCCGAGAGCTATTTTGAAGGATCTACACAATC	2640
2632	Db	GCACGGCTGCCGTGAAGTGGATGGCACCCGAGAGCTATTTTGAAGGATCTACACAATC	2691
2641	Qy	AAGAGTGACGTCTGGCTCTACGGCATCCTCTCTGGGAGATATTTTCACHTGGGTGTGAAC	2700
2692	Db	AAGAGTGACGTCTGGCTCTACGGCATCCTCTCTGGGAGATATTTTCACHTGGGTGTGAAC	2751
2701	Qy	CCTTACCTGGCAATTCCTGTCGACGCTAACTCTATAAACTGATTCAGAGTGGAATTTAAA	2760
2752	Db	CCTTACCTGGCAATTCCTGTCGACGCTAACTCTATAAACTGATTCAGAGTGGAATTTAAA	2811
2761	Qy	ATGGAGCAGCATTCATATGCCACAGAAGGATATACTTTGTAATGCAATCCTGCTGGGCT	2820
2812	Db	ATGGAGCAGCATTCATATGCCACAGAAGGATATACTTTGTAATGCAATCCTGCTGGGCT	2871
2821	Qy	TTTGACTCAAGGAAGGGGCCATCCTTCCCCAACTGACTTCATTTTAGGATGTACAGCTG	2880
2872	Db	TTTGACTCAAGGAAGGGGCCATCCTTCCCCAACTGACTTCATTTTAGGATGTACAGCTG	2931
2881	Qy	GCAGAGCGAAGAAGG-----ATGTATCAGAAACATCCATCC	2917
2932	Db	GCAGAGCGAAGAAGGCGATGTATCAGAAACATGGTGGCAACGTCCCGAAACATCCATCC	2991
2918	Qy	ATCTACCAAACAGGGGCCCTCAGCAGAGAGCGGGCTCAGAGCC-CAGTCGCCACAG	2976
2992	Db	ATCTACCAAACAGGGGCCCTCAGCAGAGAGCGGGCTCAGAGCGGCCATCGCCACAG	3051
2977	Qy	GCCAGGTGAAGATTACAGAGAAAGATTACGGAGGAGCGCTTGGACCCCGCCACCCT	3036
3052	Db	GCCAGGTGAAGATTACAGAGAAAGATTACGGAGGAGCGCTTGGACCCCGCCACCCT	3111
3037	Qy	AGCAGGCTGTAGACCGCAGAGCAAGATTAGCCTCGCCTCTGAGGAAGCGCCTACAGCG	3096
3112	Db	AGCAGGCTGTAGACCGCAGAGCAAGATTAGCCTCGCCTCTGAGGAAGCGCCTACAGCG	3171
3097	Qy	CGTTGCTTCGCTGGACTTTTCTCTAGATGCTGTGCCATTACTTCCAAAGTGACTTCAT	3156
3172	Db	CGTTGCTTCGCTGGACTTTTCTCTAGATGCTGTGCCATTACTTCCAAAGTGACTTCAT	3231
3157	Qy	AAATCAAACTCTCTCGCACAGCGGGGAGAGCCAAATATAGACTTTGTTGGTAGCCCC	3216
3232	Db	AAATCAAACTCTCTCGCACAGCGGGGAGAGCCAAATATAGACTTTGTTGGTAGCCCC	3291
3217	Qy	GCCTACCTGGGGCCCTTT-----CCACGAGCTTGAGGGGAAGCCCATGTATCTGAA	3268
3292	Db	GCCTACCTGGGGCCCTTTCCAGGCCGCCCGAGCTTGAGGGGAAGCCCATGTATCTGAA	3351
3269	Qy	ATATAGTATATCTTGTAAATACGTGAACAAACAAACCCGTTTTTTTGTCTAAGGGAAAG	3328
3352	Db	ATATAGTATATCTTGTAAATACGTGAACAAACAAACCCGTTTTTTTGTCTAAGGGAAAG	3411
3329	Qy	CTAAATATGATTTTAAAAATCTATGTTTTTAAAAATCTATGTAACTTTTTTCATCTATTTA	3388
3412	Db	CTAAATATGATTTTAAAAATCTATGTTTTTAAAAATCTATGTAACTTTTTTCATCTATTTA	3471
3389	Qy	GTGATATATTTTATGGATGGAATAAACTTTCTACTGTAAAAAATAAAAAA	3438
3472	Db	GTGATATATTTTATGGATGGAATAAACTTTCTACTGTAAAAAATAAAAAA	3521

Db	352	ACCTGCAAGTGCAGCTCGCCACCCAGGGACCTTTCCTGCTCTGGGTCCTTTAAAGCAC	411
Qy	361	AGTCCCTGGGCTGCCAGCCGCACTTTGATTTACAAAACAGAGAAATCGTTTCCATGGCC	420
Db	412	AGTCCCTGGGCTGCCAGCCGCACTTTGATTTACAAAACAGAGAAATCGTTTCCATGGCC	471
Qy	421	ATCTTGAACGTGCAGAGAGCCAGGAGGAGAAATACCTACTCCATATTCAGAGCGAAGCC	480
Db	472	ATCTTGAACGTGCAGAGAGCCAGGAGGAGAAATACCTACTCCATATTCAGAGCGAAGCC	531
Qy	481	GCCAACTACACAGTACTGTTTCACAGTCAATGTTAAGAGATACACAGCTGTATGCTAAGG	540
Db	532	GCCAACTACACAGTACTGTTTCACAGTCAATGTTAAGAGATACACAGCTGTATGCTAAGG	591
Qy	541	AGACCTTACTTTAGGAAGATGGAAAACAGGATGCATGCTCTGCATCTCCGAGGGTGT	600
Db	592	AGACCTTACTTTAGGAAGATGGAAAACAGGATGCATGCTCTGCATCTCCGAGGGTGT	651
Qy	601	CCGAGAGCCACTGTGGAGTGGTGCTCTGCAGCTCCACAGGGAAGAGCTGTAAAGAAGAA	660
Db	652	CCGAGAGCCACTGTGGAGTGGTGCTCTGCAGCTCCACAGGGAAGAGCTGTAAAGAAGAA	711
Qy	661	GGCCCTGCTGTTGCAGAAAGGAGGAAAAAGTACTTTCATGAGTTGTCGGAACAGACATC	720
Db	712	GGCCCTGCTGTTGCAGAAAGGAGGAAAAAGTACTTTCATGAGTTGTCGGAACAGACATC	771
Qy	721	AGATGCTGTGCTAGAAATGCATGGGCGCGAATGCACCAAGCTGTTCCACATAGATCTA	780
Db	772	AGATGCTGTGCTAGAAATGCATGGGCGCGAATGCACCAAGCTGTTCCACATAGATCTA	831
Qy	781	AACGAGCTCCTCAGAGCACACTSCCCAGTATTTCCTGAAAGTGGGGGAACCTTGTGG	840
Db	832	AACGAGCTCCTCAGAGCACACTSCCCAGTATTTCCTGAAAGTGGGGGAACCTTGTGG	891
Qy	841	ATCAGTGTAAAGCCATCCATGTGAACCATGGATTCGGGCTCACTCGGAGCTTGGGAAGAC	900
Db	892	ATCAGTGTAAAGCCATCCATGTGAACCATGGATTCGGGCTCACTCGGAGCTTGGGAAGAC	951
Qy	901	AAAGCCCTGGAGGAGGCGAGTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG	960
Db	952	AAAGCCCTGGAGGAGGCGAGTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG	1011
Qy	961	ATTCGGATCTCTTGGCCTTTGTGTCTTCGCTGGGAAGAACGACACCGGATATTACACC	1020
Db	1012	ATTCGGATCTCTTGGCCTTTGTGTCTTCGCTGGGAAGAACGACACCGGATATTACACC	1071
Qy	1021	TGCTCTTCTCAAAGCACCCCGCCAGTCAGCGTTGTGTGACCATCTCTAGAAAAGGGTTT	1080
Db	1072	TGCTCTTCTCAAAGCACCCCGCCAGTCAGCGTTGTGTGACCATCTCTAGAAAAGGGTTT	1131
Qy	1081	ATAAACGCTACCAGCTCCGAAGAAGAGTATGAAATTGACCCCGTACGAAAAAGTTCTCGCTC	1140
Db	1132	ATAAACGCTACCAGCTCCGAAGAAGAGTATGAAATTGACCCCGTACGAAAAAGTTCTCGCTC	1191
Qy	1141	TCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGACGCGGATCTTCTCTCAAGCCTCA	1200
Db	1192	TCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGACGCGGATCTTCTCTCAAGCCTCA	1251
Qy	1201	TTTCCTTGTGAACAGAGAGGCGCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT	1260
Db	1252	TTTCCTTGTGAACAGAGAGGCGCTGGAGGATGGGTACAGCATATCTAAATTTTGGGATCAT	1311
Qy	1261	AAGAACAGCCAGGAGAGTACATATTCATGTCAGAAAAATGATGAGCCCGCTTCCACAAA	1320
Db	1312	AAGAACAGCCAGGAGAGTACATATTCATGTCAGAAAAATGATGAGCCCGCTTCCACAAA	1371
Qy	1321	ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG	1380
Db	1372	ATGTTACGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAATGCTCAGCCAGCCAG	1431
Qy	1381	CGGTCCTGTTCCCTGTATGGCTACCCGCTACCCCTTGGACCTTGGAAAGAGTGTTCGGAC	1440

```
2521 TGTGACTTTGGACTGCGCCGAGACATCTGAGCGACTCAGTCTGTCAGGGGCAAC 2580
2572 TGTGACTTTGGACTGCGCCGAGACATCTGAGCGACTCAGTCTGTCAGGGGCAAC 2631
2581 GCACGGCTGCGGCTGAAGTGGATGCGACCCGAGAGCTTATTGAAGGGATCTACACAATC 2640
2632 GCACGGCTGCGGCTGAAGTGGATGCGACCCGAGAGCTTATTGAAGGGATCTACACAATC 2691
2641 AAGAGTGAGCTGTGCTCTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAC 2700
2692 AAGAGTGAGCTGTGCTCTACGGCATCTCTCTGGGAGATATTTTCACTGGGTGTGAC 2751
2701 CCTTACCCTGGCATCTCTGTCACGCTAACTTCTATAAAGTATGATGAGTGGATTAAA 2760
2752 CCTTACCCTGGCATCTCTGTCACGCTAACTTCTATAAAGTATGATGAGTGGATTAAA 2811
2761 ATGGAGAGCCATCTATGCCACAGAGGATATCTTTGTAATGCAATCTCTGGGCT 2820
2812 ATGGAGAGCCATCTATGCCACAGAGGATATCTTTGTAATGCAATCTCTGGGCT 2871
2821 TTTGACTCAAGGAAGGGCCATCTTCCCAACCTGACTTCTATTTAGGATGTCAGCTG 2880
2872 TTTGACTCAAGGAAGGGCCATCTTCCCAACCTGACTTCTATTTAGGATGTCAGCTG 2931
2881 GCAGAGGCGAGAAGAAGC-----ATGTCAGAAATCATCATCC 2917
2932 GCAGAGGCGAGAAGAAGCATGTATCAGAACATGGGTGGCAAGCTCCAGAACATCATCC 2991
2918 ATCTACCAAAACAGCGCGCCCTCAGCAGAGAGGGGCTCAGAGCC-CAGTCGCGACAG 2976
2992 ATCTACCAAAACAGCGCGCCCTCAGCAGAGAGGGGCTCAGAGCGCCATCGCCACAG 3051
2977 CCGCAGGTGAAGATTCACAGAGAAGAAGTTAGCCAGGAGGCTTGGACCCGCCACCT 3036
3052 GCCCAGGTGAAGATTCACAGAGAAGAAGTTAGCCAGGAGGCTTGGACCCGCCACCT 3111
3037 AGCAGGCTGACAGCGCAGCAAGATTAGCTCGCTCTGAGGAAGGCGCCCTCAGAGCG 3096
3112 AGCAGGCTGACAGCGCAGCAAGATTAGCTCGCTCTGAGGAAGGCGCCCTCAGAGCG 3171
3097 CGTGTCTCGCTGGACTTTCTCTAGATGCTGTCTGCCATTTACTCCAAAGTGAATCTAT 3156
3172 CGTGTCTCGCTGGACTTTCTCTAGATGCTGTCTGCCATTTACTCCAAAGTGAATCTAT 3231
3157 AAAATCAAACTCTCTCCACAGGGGGAGAGCCCAATATGAGACTTTGTTGGTGAGCCC 3216
3232 AAAATCAAACTCTCTCCACAGGGGGAGAGCCCAATATGAGACTTTGTTGGTGAGCCC 3291
3217 GCCTACCCCTGGGGCCCTTT-----CCACGAGCTTGGAGGAAAGCCATGATCTGAA 3268
3292 GCCTACCCCTGGGGCCCTTTCCAGGCCCCCGCCAGGCTTGGAGGAAAGCCATGATCTGAA 3351
3269 ATATAGTATATCTTCTTAATAGCTGAACAAACAAACCCGTTTGTGTAAGGAAAG 3328
3352 ATATAGTATATCTTCTTAATAGCTGAACAAACAAACCCGTTTGTGTAAGGAAAG 3411
3329 CTAATATATATTTTAAAAATCTATGTTTAAATATCTATGTAACCTTTTTCATCTATTTA 3388
3412 CTAATATATATTTTAAAAATCTATGTTTAAATATCTATGTAACCTTTTTCATCTATTTA 3471
3389 GTGATATATTTTATGATGAATAAATCTTCTACTGTAAATAAAAAA 3438
3472 GTGATATATTTTATGATGAATAAATCTTCTACTGTAGAAAAA 3521
```

## RESULT 13

PCF-US95-03718-1

```
; Sequence 1, Application PC/TUS9503718
; GENERAL INFORMATION:
; APPLICANT: GENENTECH, INC.
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THE
; NUMBER OF SEQUENCES: 4
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/03718
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wendy M. Lee
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 879PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 3521 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCF-US95-03718-1
```

Query Match 96.8%; Score 3344; DB 5; Length 3521;

Best Local Similarity 98.5%; Pred. No. 0;

Matches 3418; Conservative 0; Mismatches 20; Indels 32; Gaps 3;

QY	1	GGGGCTGCTACCGCGCGCTCCGAGGCGCATCGGGCGTTGGCGCAGCGACGCGACCGG	60
Db	52	GGGGCTGCTACCGCGCGCTCCGAGGCGCATCGGGCGTTGGCGCAGCGACGCGACCGG	111
QY	61	CGGCTGCTGCTGCTGTTGTTTGTCTAGTAATGATCTTGAGACCGTTACAACCAAGAC	120
Db	112	CGGCTGCTGCTGCTGTTGTTTGTCTAGTAATGATCTTGAGACCGTTACAACCAAGAC	171
QY	121	CTGCTGCTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGCTCATCAGCGGGAAG	180
Db	172	CTGCTGCTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGCTCATCAGCGGGAAG	231
QY	181	CCATCATCTGACCAATGGTGGCAGGATCCCGAGAGACCTCCAGTCTACCCCGAGGCG	240
Db	232	CCATCATCTGACCAATGGTGGCAGGATCCCGAGAGACCTCCAGTCTACCCCGAGGCG	291
QY	241	CAGAGTGAAGGAGCGGTATATGAAGCGGCCACCGTGGAGGTGGCGGAGTCTGGGTCCATC	300
Db	292	CAGAGTGAAGGAGCGGTATATGAAGCGGCCACCGTGGAGGTGGCGGAGTCTGGGTCCATC	351
QY	301	ACCTCTCAAGTGCAGCTCGCCACCCAGGGGAGCTTTCTGCTGCTCTGGGTCTTTAAGCAC	360
Db	352	ACCTCTCAAGTGCAGCTCGCCACCCAGGGGAGCTTTCTGCTGCTCTGGGTCTTTAAGCAC	411
QY	361	AGCTCCCTGGGCTGCCAGCGCGACTTTGATTATACAAAACAGAGAAATCGTTTCCATGCC	420
Db	412	AGCTCCCTGGGCTGCCAGCGCGACTTTGATTATACAAAACAGAGAAATCGTTTCCATGCC	471
QY	421	ATCTTGAACGTGACAGAGACCCGAGGAGAAATACCTACTCCATATTACAGAGCAAGCG	480
Db	472	ATCTTGAACGTGACAGAGACCCGAGGAGAAATACCTACTCCATATTACAGAGCAAGCG	531
QY	481	GCCAACTACACAGTACTGTTTCACAGTGAATGTAAGAGATACACAGCTGTATGTGCTAAG	540

Db 532 GCCAACTACACAGTACTGTTTCACAGTGAATGTAAGAGATACACAGCTGTATGCTAAGG 591  
Qy 541 AGACCTTACTTTAGGAAGATGAAACACAGAGTACTGCTCTGCATCTCCAGGGTGT 600  
Db 592 AGACCTTACTTTAGGAAGATGAAACACAGAGTACTGCTCTGCATCTCCAGGGTGT 651  
Qy 601 CCGGAGCCACTGTGGAGTGGTGTCTGTCAGCTCCACAGGAAAGCTGTAAAGAA 660  
Db 652 CCGGAGCCACTGTGGAGTGGTGTCTGTCAGCTCCACAGGAAAGCTGTAAAGAA 711  
Qy 661 GGCCTGTCTGTGTCAGAAAGAGGAAAGTACTTTCATGAGTGTTCGGAACAGACATC 720  
Db 712 GGCCTGTCTGTGTCAGAAAGAGGAAAGTACTTTCATGAGTGTTCGGAACAGACATC 771  
Qy 721 AGATGCTGTCTGTCAGAAAGAGGAAAGTACTTTCATGAGTGTTCGGAACAGACATC 780  
Db 772 AGATGCTGTCTGTCAGAAAGAGGAAAGTACTTTCATGAGTGTTCGGAACAGACATC 831  
Qy 781 AACCAAGGCTCTCAGAGCACACTGCCCCAGTTTATTCCTGAAAGTGGGGAACCCCTTGTGG 840  
Db 832 AACCAAGGCTCTCAGAGCACACTGCCCCAGTTTATTCCTGAAAGTGGGGAACCCCTTGTGG 891  
Qy 841 ATCAGGTGTAGGCCATCCATGTGAACCATGGATTGCGGCTCACCTGGAGCTGGAAGAC 900  
Db 892 ATCAGGTGTAGGCCATCCATGTGAACCATGGATTGCGGCTCACCTGGAGCTGGAAGAC 951  
Qy 901 AAAGCCCTGGAGGGAGGCTACTTTGAGATGAGTACTTCTCACAACAGGACCATG 960  
Db 952 AAAGCCCTGGAGGGAGGCTACTTTGAGATGAGTACTTCTCACAACAGGACCATG 1011  
Qy 961 ATTCGGATTCTTGGCCCTTGTGCTTCCGTTGGGAAGAACACACCGGATATTACACC 1020  
Db 1012 ATTCGGATTCTTGGCCCTTGTGCTTCCGTTGGGAAGAACACACCGGATATTACACC 1071  
Qy 1021 TGCCTCTCCTCAAGCACCCAGCCAGTCAGCGTTGGTGACCATCTAGAAAAAGGGTT 1080  
Db 1072 TGCCTCTCCTCAAGCACCCAGCCAGTCAGCGTTGGTGACCATCTAGAAAAAGGGTT 1131  
Qy 1081 ATAAAGCTACAGCTCCGCAAGAGATGAAATGGACCTAGCAAAAGTTCTGCTTC 1140  
Db 1132 ATAAAGCTACAGCTCCGCAAGAGATGAAATGGACCTAGCAAAAGTTCTGCTTC 1191  
Qy 1141 TCAGTCAGGTTTAAAGCGTACCACAGATCCGATGACGTGATCTTCTCAAGCCCTCA 1200  
Db 1192 TCAGTCAGGTTTAAAGCGTACCACAGATCCGATGACGTGATCTTCTCAAGCCCTCA 1251  
Qy 1201 TTTCTCTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1260  
Db 1252 TTTCTCTGTGAACAGAGAGGCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATCAT 1311  
Qy 1261 AAGAACAGCCAGGAGATACATATCTATGCAGAAATGATGACGCCAGTTTCAACAAA 1320  
Db 1312 AAGAACAGCCAGGAGATACATATCTATGCAGAAATGATGACGCCAGTTTCAACAAA 1371  
Qy 1321 ATGTTCAAGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1380  
Db 1372 ATGTTCAAGCTGAATATAAGAAAGAAACCTCAAGTGTAGCAAAATGCTCAGCCAGCCAG 1431  
Qy 1381 GCGTCTCTTCTCTGTATGGCTACCCGCTACCCCTCTTGGACCTGGAAGAGTGTTCGGAC 1440  
Db 1432 GCGTCTCTTCTCTGTATGGCTACCCGCTACCCCTCTTGGACCTGGAAGAGTGTTCGGAC 1491  
Qy 1441 AAATCTCCAAATTTGACAGGAGAAATCCAGAAAGAGTTTGAATATAAGAGCTTACAGA 1500  
Db 1492 AAATCTCCAAATTTGACAGGAGAAATCCAGAAAGAGTTTGAATATAAGAGCTTACAGA 1551  
Qy 1501 AAAGTGTGTCAGCTGGTCTGACAGTACTTCTAAATATACATGACGTGAGCCGGAAAGG 1560  
Db 1552 AAAGTGTGTCAGCTGGTCTGACAGTACTTCTAAATATACATGACGTGAGCCGGAAAGG 1611  
Qy 1561 CTTCCTGGTCAATGCTGCGGTACAAATCTATGGGACGCTCTGCGAAACCATCTTTTATA 1620  
Db 1612 CTTCCTGGTCAATGCTGCGGTACAAATCTATGGGACGCTCTGCGAAACCATCTTTTATA 1671

Qy 1621 AACTCACCAGGCCCCCTTCCCTTTTCATCCAAAGACAACATCTCCTTCTATGCGACCATGGG 1680  
Db 1672 AACTCACCAGGCCCCCTTCCCTTTTCATCCAAAGACAACATCTCCTTCTATGCGACCATGGG 1731  
Qy 1681 CTCTGTCTCCCTTTCATTTGTTTCTCATTGTGTTTCTGATCTGCCACAATAACAAAAAGCAA 1740  
Db 1732 CTCTGTCTCCCTTTCATTTGTTTCTCATTGTGTTTCTGATCTGCCACAATAACAAAAAGCAA 1791  
Qy 1741 TTTAGTAGCAGAGTACAGCTGCAGATGATCCAGGTGACTGGCCCCCTGGATACGAGTAC 1800  
Db 1792 TTTAGTAGCAGAGTACAGCTGCAGATGATCCAGGTGACTGGCCCCCTGGATACGAGTAC 1851  
Qy 1801 TTCTACGTTGACTTTCAGGACATATGAATATGACCTTAAAGTGGGAGTTCCCGAGAGACAAC 1860  
Db 1852 TTCTACGTTGACTTTCAGGACATATGAATATGACCTTAAAGTGGGAGTTCCCGAGAGACAAC 1911  
Qy 1861 TTAGAGTTTGGGAAGTCTTGGGCTTGGCGCTTTTCGGAGGGTGTATGAACGCCACCGCC 1920  
Db 1912 TTAGAGTTTGGGAAGTCTTGGGCTTGGCGCTTTTCGGAGGGTGTATGAACGCCACCGCC 1971  
Qy 1921 TATGGCATTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGTGTCTAAAGAGAAA 1980  
Db 1972 TATGGCATTAGTAAACGGGAGTCTCAATTCAGGTGGCGGTGAAGTGTCTAAAGAGAAA 2031  
Qy 1981 GCTGACAGCTGTCAAAAGCAAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2040  
Db 2032 GCTGACAGCTGTCAAAAGCAAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGA 2091  
Qy 2041 CACCATGACAACATCGTGAATCTGCTGGGGCATGCACACTGTCAAGGGCCAGTGTACTTG 2100  
Db 2092 CACCATGACAACATCGTGAATCTGCTGGGGCATGCACACTGTCAAGGGCCAGTGTACTTG 2151  
Qy 2101 ATTTTGAATATTTGCTGATGCTGACCTCCTCAACTACCTAAGAGTAAAGAGAGAAAG 2160  
Db 2152 ATTTTGAATATTTGCTGATGCTGACCTCCTCAACTACCTAAGAGTAAAGAGAGAAAG 2211  
Qy 2161 TTTTCACAGGACATGGACAGAGATTTTAAAGGAACATAATTTACCTTCTTACCCTACTTTC 2220  
Db 2212 TTTTCACAGGACATGGACAGAGATTTTAAAGGAACATAATTTACCTTCTTACCCTACTTTC 2271  
Qy 2221 CAGGCACATTTCAAAATCCAGCATGCTGCTGTTACAGAGAGTTTCAAGTTACACCCGCCCTTG 2280  
Db 2272 CAGGCACATTTCAAAATCCAGCATGCTGCTGTTACAGAGAGTTTCAAGTTACACCCGCCCTTG 2331  
Qy 2281 GATCAGCTCTCAGGTTTCAATGGGAATTCATTCATCTGAAATGAGATGAGATGAAATGAA 2340  
Db 2332 GATCAGCTCTCAGGTTTCAATGGGAATTCATTCATCTGAAATGAGATGAGATGAAATGAA 2391  
Qy 2341 ACCCAGAGAGCTGGCAGAAAGAGAGGAGAGATTTGAACGCTGCTGACGTTTGAAGAC 2400  
Db 2392 ACCCAGAGAGCTGGCAGAAAGAGAGGAGAGATTTGAACGCTGCTGACGTTTGAAGAC 2451  
Qy 2401 CTCTTTTGTGTTTGGCTTACCAAGTGGCAAGAGGATGAAATTTCTGGAGTTCAAGTCTGT 2460  
Db 2452 CTCTTTTGTGTTTGGCTTACCAAGTGGCAAGAGGATGAAATTTCTGGAGTTCAAGTCTGT 2511  
Qy 2461 GTCCACAGAGACTGGCAGCCAGGAATGTGTTGGTCAACCCACGGAGGTGGTGAAGATC 2520  
Db 2512 GTCCACAGAGACTGGCAGCCAGGAATGTGTTGGTCAACCCACGGAGGTGGTGAAGATC 2571  
Qy 2521 TGTGACTTTGGACTGGCCCGAGACATCTCTGAGGAGCTCCAGCTACGTCAGGGGCAAC 2580  
Db 2572 TGTGACTTTGGACTGGCCCGAGACATCTCTGAGGAGCTCCAGCTACGTCAGGGGCAAC 2631  
Qy 2581 GCACGGCTCCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACACAATC 2640  
Db 2632 GCACGGCTCCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGATCTACACAATC 2691  
Qy 2641 AAGAGTGAAGTGTCTTACGGCATCTCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2700  
Db 2692 AAGAGTGAAGTGTCTTACGGCATCTCTCTCTGGGAGATATTTTCACTGGGTGTGAAC 2751



QY 2701 CCTTACCTGGCATTCCTGTGACGCTAACTTCTATAACTGATTCAGAGTGATTTAA 2760  
Db 2752 CCTTACCTGGCATTCCTGTGACGCTAACTTCTATAACTGATTCAGAGTGATTTAA 2811  
QY 2761 ATGGAGCAGCCATTCATGCCACAGAGGATATCTTTGTAATGCAATCCTGCGGCT 2820  
Db 2812 ATGGAGCAGCCATTCATGCCACAGAGGATATCTTTGTAATGCAATCCTGCGGCT 2871  
QY 2821 TTTGACTCAAGGAAGCGGCATCTTCCCACTGACTTCAATTTAGGATGTCAGTG 2880  
Db 2872 TTTGACTCAAGGAAGCGGCATCTTCCCACTGACTTCAATTTAGGATGTCAGTG 2931  
QY 2881 CGAGAGGAGAGAGC-----ATGTATCAGAAATCATCATCC 2917  
Db 2932 CGAGAGGAGAGAGAGCGATGTATCAGAACATGGTGGCAAGCTCCAGAAATCATCATCC 2991  
QY 2918 ATCTACCAAAACAGCGGCGCTCAGCAGAGAGGCGGCTCAGAGCC-CAGTGGCCACAG 2976  
Db 2992 ATCTACCAAAACAGCGGCGCTCAGCAGAGAGGCGGCTCAGAGCCGCGCATCGCCACAG 3051  
QY 2977 GCCCAGGTCAGATTCACAGAGAAAGAGTTAGCGCTTGGAGGAGGCGCTTGGACCCGCCACCT 3036  
Db 3052 GCCCAGGTCAGATTCACAGAGAAAGAGTTAGCGAGGAGGCGCTTGGACCCGCCACCT 3111  
QY 3037 AGCAGGCTGTAGACGCGAGAGCAAGATTAGCCTCGCCTCTGAGGAAGCGCCCTACAGCG 3096  
Db 3112 AGCAGGCTGTAGACGCGAGAGCAAGATTAGCCTCGCCTCTGAGGAAGCGCCCTACAGCG 3171  
QY 3097 CGTTGCTTGGCTGAGCTTTCTCTAGATGCTCTGCTGCCATTAATCTCCAAAGTACTCTAT 3156  
Db 3172 CGTTGCTTGGCTGAGCTTTCTCTAGATGCTCTGCTGCCATTAATCTCCAAAGTACTCTAT 3231  
QY 3157 AAATCAAACTCTCTGCGCAGAGGCGGAGAGCAATTAATGAGACTTTGTTGGTGGCC 3216  
Db 3232 AAATCAAACTCTCTGCGCAGAGGCGGAGAGCAATTAATGAGACTTTGTTGGTGGCC 3291  
QY 3217 GCCTACCTCGGGGCGCTT-----CCACGAGCTTGGAGGAGGAGCCATGATCTGAA 3268  
Db 3292 GCCTACCTCGGGGCGCTTCCAGGCGCCCGCAGGCTTGGAGGAGGAGCCATGATCTGAA 3351  
QY 3269 APATAGTATATTTGTAATACGTGAACAAACCAACCCCGTTTTTCTAAGGGAAG 3328  
Db 3352 ATATAGTATATTTGTAATACGTGAACAAACCAACCCCGTTTTTCTAAGGGAAG 3411  
QY 3329 CTAATATGATTTTTAAATCTATGTTTTAAATCTATGTAACCTTTTCTATCTATTTA 3388  
Db 3412 CTAATATGATTTTTAAATCTATGTTTTAAATCTATGTAACCTTTTCTATCTATTTA 3471  
QY 3389 GTGATATATTTATGATGGAATAAATCTTCTACTGTAAAAAATAAAAA 3438  
Db 3472 GTGATATATTTATGATGGAATAAATCTTCTACTGTAAAAAATAAAAA 3521

RESULT 14  
US-07-977-451-3

; Sequence 3, Application US/07977451  
; Patent No. 5270458  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/977,451  
; FILING DATE: 19921119  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US UNASSIGNED  
; FILING DATE: 12-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/679,666  
; FILING DATE: 02-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Felt, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: LEM-3-7P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405  
; TELEFAX: 212-645-2054  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3501 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: CDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 58..3039  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: 139..3036  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: 58..138  
; US-07-977-451-3

Query Match 64.5%; Score 2227.2; DB 1; Length 3501;  
Best Local Similarity 79.6%; Pred. No. 0;  
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;  
QY 8 GGCTACCGCGCTCCGGAGGCCATCGGCGCTTGGCGGAGCGGAGCGGCGGCTGC 67  
Db 35 GGGGACCCCGGCGCTCCGGAGGCCATCGGCGCTTGGC---GCGGACGCGGCGCACCGTGC 91  
QY 68 TGCTGCTGTTGTTTGTCTAGTAATGATTTCTTGAGACCGTTACAAACCAAGACCTGCTG 127  
Db 92 CGTGTCTGTTGTTTCTGCAATGATATTTGGGACTATTAACAATCAAGATCTGCTG 151  
QY 128 TGATCAAGTGTGTTTAAATCAGTCATGAGAACATGCTCATCAGCGGAAAGCCATCAT 187

152 TGATCAAGTGTGTTTAAATCAATATAAGAAATGATTCATCAGTGGGGAAGTCATCAT 211  
188 CGTACCCAAATGGTCGAGGATCCCAAGAGACCTCCAGTGTACCCGAGGCGCCAGAGTG 247  
212 CATATCCCATGGTATCAGATATCCCGGAAGACCTCGGGTGTGAGACCCAGAGCT 271  
248 AAGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCGGAGTCTGGGTCCATACCCCTGC 307  
272 CAGGGACAGTGTACGAAGCTGCCGCTGTGGAAGTGGATGTATCTGCTTCCATCACACTGC 331  
308 AAGTGCAGCTGCACACCCAGGGGACCTTTCTCCCTCTCGGCTCTTAAAGCACAGCTCCC 367  
332 AAGTGCCTGGTGCATGCGCCAGGGAACATTTCTCTCTGGGTCTTTAAGCACAGCTCCC 391  
368 TGGGCTGCCAGCCGCACTTTGATTTACAAACACAGGAATCGTTTCCATGCCATCTTGA 427  
392 TGAATTTCCAGCCACATTTTGAATTTACAAACAGAGGAGTTGTTTCCATGGTCTCAATTTGA 451  
428 ACGTGACAGAGACCCAGCGAGGAATACCTTACTTCCATATTCAGAGCGAAGCGGCCAACT 487  
452 AATGACAGAAACCCAGCTGGAGAAATACCTTCTTTTTCAGAGTGAAGCTACCAAT 511  
488 ACACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGTCTTAAGGAGACCTT 547  
512 ACACAAATATTGTTTACAGTGAATGTAAGAAATACCTTCTTTTACACATTAAGAAGACCTT 571  
548 ACTTTAGAGATGGAACCCAGGATGCATGCTCTGCATCTCCGAGGCTTCCGGAGC 607  
572 ACTTTAGAAAATGGAACCCAGAGCCCTGGTCTGCATATCTGAGAGCCCTCCAGC 631  
608 CCACTGTGGAGTGGGTCTCTGAGCTCCACAGGGAAGCTGTAAAGAAAGAGCCCTG 667  
632 CGATCTGGAAATGGGTGCTTGGCAATTCACAGGGGGAAGCTGTAAAGAAAGTCCAG 691  
668 CTGTTGTGAGAAAGGAGGAAAGTACTTTCATGAGTTGTTGGAACACATCAGATGCT 727  
692 CTGTTGTTAAAGAGGAGGAAAGTGTCTCATGAATTTTGGGACGACATAAGGTCT 751  
728 GTGCTAGAAATGCACTGGGCGGCAATGCACCAAGCTGTTCCACCATAGATTAACACAG 787  
752 GTGCCAGAAATGAATCGGAGGGAATGCACAGGCTGTTCCACATAGATTAATCAAA 811  
788 CTCTCAGAGCACACTGCCAGTATTCTCTGAAGTGGGGAACCTTGTGGATCAGGT 847  
812 CTCTCAGACCAATGCCACAATTTCTTAAAGTAGGGGAACCTTATGGATAAGGT 871  
848 GTAAGGCCATCCATGTGAACCATGGATTGCGGCTCAGCTGGAGCTGGAAGACAAAGCC 907  
872 GCAAAGCTGTGATGTGAACCATGATTCGGGCTCACCTGGGAATTAGAAACAAAGCAC 931  
908 TGGAGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAAAGAGACCATGATTCGGA 967  
932 TCGAGAGGGCAACTACTTTGAGATGAGTACCTATTCAACAAACAGAACTATGATACGGA 991  
968 TTCTCTGGCTTGTCTCTCCCTGGGGAAGGACAGACCGGATATTACACCTGCTCTT 1027  
992 TTCTGTTGCTTTGTATCATCAGTGGCAAGAAACGACACCGGATACTACACTTGTTCCT 1051  
1028 CCTCAAGACGCCAGCGAGTCAGCGTTGGTGAACCATCTCTAGAAAAGGGTTTATAACG 1087  
1052 CTTCAAGACATCCAGTCATCAGCTTTGGTTACCATCGTAGGAAGGGATTTATAATG 1111  
1088 CTACCACTCGCAGAGAGATGATAATGTACCCGTACGAAAGTCTCTCTCAGTCA 1147  
1112 CTACCAATTCAGTGAAGATTATGAATTTGACCAATATGAAGAGTTTGTGTTTTCTGTCA 1171  
1148 GGTTTAAAGCGTACCCAGCAATCCGATGCAGTGGATCTCTCAAGCCCTCATTTCTT 1207  
1172 GGTTTAAAGCGTACCCAGCAATTCAGATGATGAGTGGACCTTCTCTCGAAAATCATTTCTT 1231  
1208 GTGAACAGAGGCGCTGGAGGATGGGTACAGCATATCTAAATTTTGGCATATGAAGACA 1267

1232 GTGAGCAAAAGGGTCTTGATAACGGATACAGCATATCCAAGTTTTCGAATCATAGACACC 1291  
1268 AGCCAGGAGAGTACATATTTCTATCAGAAAATGATGAGCCAGTTCCACAAAATGTTCA 1327  
1292 AGCCAGGAGATATATATTCATGAGAAAATGATGATGCCCAATTTACAAAATGTTCA 1351  
1328 CGCTGAATATAGAAAAGAACCTCAAGTGTCTAGCAAAATCCCTTCAGCCAGCGGCTCT 1387  
1352 CGCTGAATATAGAAAGAAACCTCAAGTGTCTGAGAGAGCATCGCAAGTCAAGCGCTCT 1411  
1388 GTTCCCTGTAGTACCCGCTACCCCTCTTGACCTGGAAGAGTGTTCGGACAATCTC 1447  
1412 GTTCTCGGATGATACCATTTACCATCTTGACCTGGAAGAGTGTTCAGACAAGTCTC 1471  
1448 CCAATTCACAGGAGAAATCCAGAGAGTGTGGAATATAAAGAGCTACACAAGAGTCT 1507  
1472 CCACTGCACAGAGAGATCACAGAGAGTCTGGAATAGAAAGCTACAGAAAAGTGT 1531  
1508 TTGGCCAGTGGGTGTCGAGCAGTACTCTAAATATAGTGAAGCGGGAAGGCTTCTGG 1567  
1532 TTGGACAGTGGGTGTCGAGCAGTACTCTAAACATGAGTGAAGCCATAAAGGTTCTCTG 1591  
1568 TCAATGCTGTGCTACAAATTTCTATGGCAGCTCTTGGCAACCATCTTTTAAATCTAC 1627  
1592 TCAAGTGTGTGCATACAAATTTCCCTTGGCAGATCTTGTGAGACGATCCTTTTAACTCTC 1651  
1628 CAGGCCCCCTTCCCTTTCATCCAAGACATCTCTCTATCGGACCATTTGGGCTCTGTC 1687  
1652 CAGGCCCCCTTCCCTTTCATCCAAGACATCTCTCTATCTATGCAAAATTTGGTGTGTC 1711  
1688 TCCCTCTCATTTGTTGTTCTCTGATCTGCCACAAATACAAAAGCAATTTAGGT 1747  
1712 TCCTCTCATTTGCTTTTAACTCTGCTAAATTTGTGCAAGTACAAAAGCAATTTAGGT 1771  
1748 ACGAGATCAGCTGCAGATGATCCAGTGTGAGTGGCCCTGGATACCGAGTACTCTACG 1807  
1772 ATGAAAGCCAGCTACAGATGATGAGTACAGGTGACCGGCTCTCTCAGATATGATCTACG 1831  
1808 TTGACTTCAGGCACTATGAATATGACCTTAAGTGGAGTTCCTCGAGAGAGCACTTAGGT 1867  
1832 TTGATTTTCAGAGAAATGAATATGATCTCAATGGAGTTCCTCAAGAGAAAATTTAGGT 1891  
1868 TTGGAAAGCTCTGGGGTCTGGCGCTTTCGGAGGTTGATGAACGCCAGCTCTATGGCA 1927  
1892 TTGGAGAGTACTAGGATCAGGTGCTTTTGGAAAAGTGTGAACGCCAGAGCTTATGGA 1951  
1928 TTACTTAAACGGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAAGAGAAAGCTGACA 1987  
1952 TTAGCAAAACAGGAGTCTCAATCCAGTTGCCGTCAAAATGCTGAAAGAAAAGCAGACA 2011  
1988 GCTGTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047  
2012 GCTCTGAAGAGAGGCACTCATGTGAGAACTCAAGATGATGATGATGATGATGATGATG 2071  
2048 ACAACATGCTGAATCTGCTGGGGGATGCACACTGTGAGGCGCAGTGTGATGATGATG 2107  
2072 AGATATTGTGAACCTGCTGGGGGCTGCACACTGTGAGAGCAATTTACTTGTGATTTG 2131  
2108 AATATTGTTGCTATGTTGACCTCTCACTACTTAAAGAGTAAAAGAGAGAGTTCACA 2167  
2132 AATACTGTTGCTATGTTGATCTCTCACTATCTAAGAGTAAAAGAGAGAAATTTACA 2191  
2168 GGACATGACAGAGATTTTAAAGAACATAATTTCAAGTCTTCTTACCTACTTTCCAGGCAC 2227  
2192 GGACTTGCACAGAGATTTTCAAGGAACACAATTTCAAGTCTTACCCACTTTTCCCAATCAC 2251  
2228 ATTCAAAATCCAGATGCTGTTTCCAGAGAGTTTCAAGTCTTACCTACTTTCCAGGCAC 2287  
2252 ATCCAAATTCAGCATGCTGTTTCAAGAGAGTTCAGATACACCCGGATCGGATCAAA 2311  
2288 TCTCAGGCTTCAATCGGAAATCAATTTCAATCTGGAAGATGAGATTTGAATTAAGAAACACA 2347  
2312 TCTCAGGCTTCAATCGGAAATTCATTTCTCTGGAAGATGAAATTAATTAAGAAACCAA 2371





1892 TTGGGAGGTACTAGGATCAGGTGCTTTTGGGAAAAGTATGATGAACGCAACGCTTATGGAA 1951  
1928 TTAGTAAACGGGAGTCTCAATTCAGTGGCGGTGAAGATGCTAAAGAGAGAAAGCTGCACA 1987  
1952 TTAGCAAAACAGAGAGTCTCAATCCAGGTTCGGGTCAAAATGCTGAAGAAAGAAAGCAGACA 2011  
1988 GCTGTGAAAAGAGTCTCTCATGTGGAGGCTCAAAATGATGACCCCACTGGGACACATG 2047  
2012 GCTCTGAAGAGAGGCACTCATGTGCAAGTCAAGATGATGATGATGATGATGATGATGATG 2071  
2048 ACAACATCGTGAATCTCTGGGGCATGCACTGTGAGGGCCAGTGTACTGTGATTTTGG 2107  
2072 AGAATATGTGAACCTCTGGGGCGTGCACACTGTGAGGACCAATTTACTGTGATTTTGG 2131  
2108 AATATCTGCTATGTGTGACCTCTCAACTACCTTAAGAGTAAAGAGAGAGTTCACACA 2167  
2132 AATACTGTGCTATGTGTGATCTCTCAACTATCTTAAGAGTAAAGAGAGAGTTCACACA 2191  
2168 GGACATGGACAGAGATTTTAAAGGAACATAATTTCACTTCTTACCCCTACTTTCCAGGCAC 2227  
2192 GGAATGGACAGAGATTTTCAAGGAACACATAATTTCACTTCTTACCCCACTTTTCCCAATCAC 2251  
2228 ATTCAATTCAGCATGCTGTTTCAGAGAGTTCAGTTCACACCGCCCTGTGATCAGC 2287  
2252 ATCCAAATTCAGCATGCTGTTTCAGAGAGTTCAGTTCACACCGCCCTGTGATCAGC 2311  
2288 TCTCAGGTTCAATGGGAATTCATCTCAAGATGAGATGAATATGAAATGAAATGAAATGAA 2347  
2312 TCTCAGGTTCAATGGGAATTCATCTCAAGATGAGATGAATATGAAATGAAATGAAATGAA 2371  
2348 AGAGCTGGGAGAGAGAGAGAGATTTGAAGCTGCTGACCTTTTGAAGACCTCCCTTT 2407  
2372 AAAGGCTG-----GAAGAGAGAGAGATTTGAAGCTGCTGACCTTTTGAAGACCTCCCTTT 2425  
2408 GCTTTGGGTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCGTGTGCCACA 2467  
2426 GCTTTGGATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCGTGTGCCACA 2485  
2468 GAGACCTGGGAGCAGAGATGTTGTTGTCACCCAGGAGGTGGTGAAGATCTGTGACT 2527  
2486 GAGACCTGGGAGCAGAGATGTTGTTGTCACCCAGGAGGTGGTGAAGATCTGTGACT 2545  
2528 TTGGACTGGCCCGGAGACATCTCTGAGGAGTCCAGCTACGTCAGGAGGCAACGACGCGC 2587  
2546 TTGGATTTGGCTCGAGATATCATGATGATGATGATGATGATGATGATGATGATGATGATG 2605  
2588 TGCCGGTGAAGTGAATGGCACCAGAGAGCTTATTTGAAGGATCTACACATCAAGAGTG 2647  
2606 TGCCGTGAAAATGGAATGGGCCCCGAAAGCCCTGTTTGAAGGATCTACACCATTAAGAGTG 2665  
2648 ACCTGCTGCTACGGCATCTCTCTCGGAGATATTTTCACTGGGTGTGAACCCCTTACC 2707  
2666 ATGCTGCTCATATGGAATATTTACGTGGGAAATCTCTCACTGGGTGTGAATCTCTACC 2725  
2708 CTGGCATTCCTGTCGACGCTAACTTCTATAAAGTGAATGATTCAGAGTGGATTTAAATGGAGC 2767  
2726 CTGGCATTCCTGTCGATCTAACTTCTATAAAGTGAATGATTCAGAGTGGATTTAAATGGAGC 2785  
2768 AGCCATTCCTGTCGACGCTAACTTCTATAAAGTGAATGATTCAGAGTGGATTTAAATGGAGC 2827  
2786 AGCCATTCCTGTCGACGCTAACTTCTATAAAGTGAATGATTCAGAGTGGATTTAAATGGAGC 2845  
2828 CAAGGAAGCGCCATCTCTCCCACTTCACTTCTAGGATGTCAGCTGGGAGG 2887  
2846 CAAGGAAGCGCCATCTCTCCCACTTCACTTCTAGGATGTCAGCTGGGAGG 2905  
2888 CAGAAGAG-----CATGTATCAGAAATCATCCATCTCTACC 2924  
2906 CAGAAGAGCGATGATCAGATGTTGGATGGCCGCTGTTTCGGAATGTCCTCACACCTACC 2965  
2925 AAACAGGCGGCCCTCAGCAGAGAGGCGGCTCAGAGC-COAGTCGCCACAGCGCCAGG 2983  
2966 AAACAGGCGGCTTTTCAAGCAGAGAGATGATTTGGGGCTACTCTCTCCGCGGCTCAGG 3025

3084 T-GAAGATTACAGAGAGAGAGAGTTCAGGAGGCGCTTGGACCCCGCC-----ACCCTAG 3038  
3026 TCGAAGATTTCGTAGAGGAACAATTTAGTTTAAAGGACTTCATCCCTCCACCTATATCCCTAA 3085  
3039 CAGGCTGTAGACCCGAGAGCAAGATTAGCTCGCTCT--GAGGAAGCGCCCTACACGC 3096  
3086 CAGGCTGTAGATTACCAAAACAAGATTAAATTTTCATCTAAAGAAAATCTATTATCAAC 3145  
3097 CGTTGCTTCGCTGACCTTTTCTCTAGATGCTGTCTGCCATTACTC-----CAAAGTGA 3149  
3146 TGCTGCTTCACACACTTTTCTCTAGAGCGCTCTGCGCTTACCTTGTCTTTTCAAAGGGA 3205  
3150 CTTCATAAAATCAAACTCTCTCGCACAGCGGGGAGAGCCAAATATGAGACTTGTGG 3209  
3206 CTTTTGTAATAATCAATCATCTCTGTCACAAGCGAGGAGCTGATAATGAACTTTATTG 3265  
3210 TGACCCCGCTTACCTCGGGGCTTTCCACGAGCTTGAGGGAAAGCCATGTATCTGAAA 3269  
3266 GAGCATTGATCTGATCCAGCGCTTCTCAGCGCGCTTGAGTGAATTTGTGATCTGAAG 3325  
3270 TATAGTATATTCTGTAAATACGTGAACAAACAAACCGCTTTTTCCTAAGGGAAGC 3329  
3326 TACAGTATATTCTGTAAATACATAAAACAAA-----ACCATTTTCTAAGGAGAGC 3378  
3330 TAAATATGATTTTAAATAATCTATGTTTAAATAATCTATGTAACCTTTTTCATCTATTAG 3389  
3379 TAATATGATTTT--AAGTCTATCTTTTAAATAATCTATGTAACCTTTTTCAGCTATTAG 3436  
3390 TGATATATTTTATGATGGAATAAATCTTCTACTGTAAAAAATAAATAAATAAATAAATAA 3449  
3437 TGATATATTTTATGAGTGGGAATAAATTTCTACTACAGAAAAAATAAATAAATAAATAA 3496  
3450 AAAA 3453  
3497 AAAA 3500

Search completed: May 25, 2003, 12:20:02  
Job time : 153.496 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 11:51:23 ; Search time 398.232 Seconds  
(without alignments)  
11449.495 Million cell updates/sec

Title: US-09-919-408-1

Perfect score: 3453

Sequence: 1 GCGGCTGGTACCGCGCC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published\_Applications\_NA.:

```
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3453	100.0	3453	10	US-09-919-408-1
2	3453	100.0	3453	10	US-09-872-136-1
3	2227.2	64.5	3501	10	US-09-919-408-3
4	2227.2	64.5	3501	10	US-09-872-136-3
5	2081.8	60.3	3120	10	US-09-982-610-22
6	466.6	13.5	645	9	US-09-796-692-7538
7	466.6	13.5	645	9	US-10-040-862-7538
8	456.4	13.2	610	9	US-09-796-692-8085
9	456.4	13.2	610	9	US-10-040-862-8085
10	454.6	13.2	605	9	US-09-796-692-8927
11	454.6	13.2	605	9	US-10-040-862-8927
12	448.8	13.0	597	9	US-09-796-692-7976
13	448.8	13.0	597	9	US-10-040-862-7976
14	448	13.0	601	9	US-09-796-692-8273
15	448	13.0	601	9	US-10-040-862-8273
16	448	13.0	608	9	US-09-796-692-8132
17	448	13.0	608	9	US-10-040-862-8132
18	447.2	13.0	597	9	US-09-796-692-9010
19	447.2	13.0	597	9	US-10-040-862-9010

20	440	12.7	605	9	US-09-796-692-8047	Sequence 8047, Ap
21	440	12.7	605	9	US-10-040-862-8047	Sequence 8047, Ap
22	439.8	12.7	589	9	US-09-796-692-8630	Sequence 8630, Ap
23	439.8	12.7	589	9	US-10-040-862-8630	Sequence 8630, Ap
24	424.4	12.3	594	9	US-09-796-692-8135	Sequence 8135, Ap
25	424.4	12.3	594	9	US-10-040-862-8135	Sequence 8135, Ap
26	423.2	12.3	564	9	US-09-796-692-9025	Sequence 9025, Ap
27	423.2	12.3	564	9	US-10-040-862-9025	Sequence 9025, Ap
28	416.8	12.1	556	9	US-09-796-692-7883	Sequence 7883, Ap
29	416.8	12.1	556	9	US-10-040-862-7883	Sequence 7883, Ap
30	407.2	11.8	550	9	US-09-796-692-8494	Sequence 8494, Ap
31	407.2	11.8	550	9	US-10-040-862-8494	Sequence 8494, Ap
32	399.8	11.6	539	9	US-09-796-692-8255	Sequence 8255, Ap
33	399.8	11.6	539	9	US-10-040-862-8255	Sequence 8255, Ap
34	386.8	11.2	510	9	US-09-796-692-8268	Sequence 8268, Ap
35	386.8	11.2	510	9	US-10-040-862-8268	Sequence 8268, Ap
36	360.2	10.4	497	9	US-09-796-692-8408	Sequence 8408, Ap
37	360.2	10.4	497	9	US-10-040-862-8408	Sequence 8408, Ap
38	357.4	10.4	3992	10	US-09-944-807-9	Sequence 9, Appli
39	350	10.1	606	9	US-09-796-692-8114	Sequence 8114, Ap
40	350	10.1	606	9	US-10-040-862-8114	Sequence 8114, Ap
41	342.2	9.9	508	9	US-09-796-692-8032	Sequence 8032, Ap
42	342.2	9.9	508	9	US-10-040-862-8032	Sequence 8032, Ap
43	329	9.5	5084	9	US-09-954-531-1383	Sequence 1383, Ap
44	329	9.5	5084	9	US-10-192-867-3	Sequence 3, Appli
45	329	9.5	5084	10	US-09-967-768A-277	Sequence 277, App

#### ALIGNMENTS

RESULT 1  
US-09-919-408-1  
Sequence 1, Application US/09919408  
Patent No. US20020072077A1  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL RECEPTORS AND THEIR LIGANDS

NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ImClone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/919,408

FILING DATE: 31-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/977,451

FILING DATE: <Unknown>

APPLICATION NUMBER: US 07/906,397

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: US PCT/US92/05401

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: TW 81102961

FILING DATE: 15-APR-1992

APPLICATION NUMBER: US PCT/US92/02750

FILING DATE: 02-APR-1992

APPLICATION NUMBER: US 07/813,593

FILING DATE: 24-DEC-1991

APPLICATION NUMBER: US 07/793,065

FILING DATE: 15-NOV-1991

APPLICATION NUMBER: US 07/728,913

FILING DATE: 28-JUN-1991





QY 1681 CTCGTCTCCCTTCATTGTTGTTCTCATTTGTTGATCTGCCCAAAATACAAAAGCAA 1740  
DB CTCGTCTCCCTTCATTGTTGTTCTCATTTGTTGATCTGCCCAAAATACAAAAGCAA 1740  
QY 1741 TTTAGTACGAGAGTACAGTACAGATGATCAGGTGACTGGCCCTCGATTAACAGGTAC 1800  
DB TTTAGTACGAGAGTACAGTACAGATGATCAGGTGACTGGCCCTCGATTAACAGGTAC 1800  
QY 1801 TTCTAGTTGACTTCAGGACTATGAATATGACCTTAAGTGGAGTTCCTCCGAGAGAAC 1860  
DB TTCTAGTTGACTTCAGGACTATGAATATGACCTTAAGTGGAGTTCCTCCGAGAGAAC 1860  
QY 1861 TTAGAGTTTGGAGGCTCTGGGCTCGCGCTTTTCGGAGGGTGAACCCACGGCC 1920  
DB TTAGAGTTTGGAGGCTCTGGGCTCGCGCTTTTCGGAGGGTGAACCCACGGCC 1920  
QY 1921 TATGCAATTAGTAAACGGGAGTCTCAATTTCAGGTGGGGTGAAGATGCTAAAGAGAAA 1980  
DB TATGCAATTAGTAAACGGGAGTCTCAATTTCAGGTGGGGTGAAGATGCTAAAGAGAAA 1980  
QY 1981 GCTGACAGCTGTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040  
DB GCTGACAGCTGTGAAAAGAGCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGA 2040  
QY 2041 CACCATGACACATGCTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTGTACTTG 2100  
DB CACCATGACACATGCTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTGTACTTG 2100  
QY 2101 ATTTTGAATATTTGCTATGCTGACCTCCTCAACTACCTAAGAGTAAAGAGAGAG 2160  
DB ATTTTGAATATTTGCTATGCTGACCTCCTCAACTACCTAAGAGTAAAGAGAGAG 2160  
QY 2161 TTTACAGGACATGGACAGAGATTTTAAAGAACATAATTCAGTTCTTACCCCTACTTTC 2220  
DB TTTACAGGACATGGACAGAGATTTTAAAGAACATAATTCAGTTCTTACCCCTACTTTC 2220  
QY 2221 CAGGACATTCAAATTCAGAGTCCCTGTTACAGAGAGTTCAGTTACACCCGCCCTTG 2280  
DB CAGGACATTCAAATTCAGAGTCCCTGTTACAGAGAGTTCAGTTACACCCGCCCTTG 2280  
QY 2281 GATCAGCTCTCAGGTTCAATGGGAATTCATTTCAATTCGAAGATGAGATGAATGAA 2340  
DB GATCAGCTCTCAGGTTCAATGGGAATTCATTTCAATTCGAAGATGAGATGAATGAA 2340  
QY 2341 AACGAGAGAGGCTGGCAGAGAGAGAGAGAGATTTGAACGCTGCTGACGTTTGAAGAC 2400  
DB AACGAGAGAGGCTGGCAGAGAGAGAGAGAGATTTGAACGCTGCTGACGTTTGAAGAC 2400  
QY 2401 CTCCTTTGCTTGGTACCAAGTGGCAAGGATGGAATTCCTGGAGTTCAAGTCTGT 2460  
DB CTCCTTTGCTTGGTACCAAGTGGCAAGGATGGAATTCCTGGAGTTCAAGTCTGT 2460  
QY 2461 GTCCACAGAGACTGGCAGCAGGAATGTTGGTCACCCAGGGAAGTGTGAAGATC 2520  
DB GTCCACAGAGACTGGCAGCAGGAATGTTGGTCACCCAGGGAAGTGTGAAGATC 2520  
QY 2521 TGTGACTTTGGACTGGCCGAGACATCCTGAGCGACTCCAGTACGTCGTGAGGGCAAC 2580  
DB TGTGACTTTGGACTGGCCGAGACATCCTGAGCGACTCCAGTACGTCGTGAGGGCAAC 2580  
QY 2581 GCACGGCTGCCGGTGAAGTGAATGATGACCCAGAGCTTATTTGAAGGGATCTACAAATC 2640  
DB GCACGGCTGCCGGTGAAGTGAATGATGACCCAGAGCTTATTTGAAGGGATCTACAAATC 2640  
QY 2641 AAGAGTACGCTGCTGCTAGGCTCTCTCTGGAGATATTTTCACTGGGTGAGC 2700  
DB AAGAGTACGCTGCTGCTAGGCTCTCTCTGGAGATATTTTCACTGGGTGAGC 2700  
QY 2701 CCTTACCCTGGCATTCCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGGATTTAA 2760  
DB CCTTACCCTGGCATTCCTGTCGACGCTAACTTCTATAAACTGATTCAGAGTGGATTTAA 2760

QY 2761 ATGGAGCAGCCATTCTATGCCACAGAGGATATACCTTTGTAATGCAATCCTCTGGGCT 2820  
DB ATGGAGCAGCCATTCTATGCCACAGAGGATATACCTTTGTAATGCAATCCTCTGGGCT 2820  
QY 2821 TTTGACTCAAGGAAGCGGCATCTTCCCACACCTGACTTCATTTTAGGATGTCAGCTG 2880  
DB TTTGACTCAAGGAAGCGGCATCTTCCCACACCTGACTTCATTTTAGGATGTCAGCTG 2880  
QY 2881 GCAGAGCAGAGAAGATGATCAGAACATCCATCATCTACCAACAGCGGCGCCCT 2940  
DB GCAGAGCAGAGAAGATGATCAGAACATCCATCATCTACCAACAGCGGCGCCCT 2940  
QY 2941 CAGCAGAGAGCGGCTCAGAGCCAGCTGCCACAGCGCAGGTGAAGATTCACAGAA 3000  
DB CAGCAGAGAGCGGCTCAGAGCCAGCTGCCACAGCGCAGGTGAAGATTCACAGAA 3000  
QY 3001 AGAAGTTAGCAGAGGCGCTTGAGCCCGCCCTAGCAGGCTGTAGACCGCAGAGCCA 3060  
DB AGAAGTTAGCAGAGGCGCTTGAGCCCGCCCTAGCAGGCTGTAGACCGCAGAGCCA 3060  
QY 3061 AGATTAGCCTCCCTCTGAGGAAGCGCTACAGCGCTTGCCTGCGTGGACTTTTCTCT 3120  
DB AGATTAGCCTCCCTCTGAGGAAGCGCTACAGCGCTTGCCTGCGTGGACTTTTCTCT 3120  
QY 3121 AGATGCTGCTGCCATTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCGCACAG 3180  
DB AGATGCTGCTGCCATTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCGCACAG 3180  
QY 3181 GCGGAGAGCCCAATATGAGACTTGTGTGAGCGCCCTACCTGGGCGCTTTCCACG 3240  
DB GCGGAGAGCCCAATATGAGACTTGTGTGAGCGCCCTACCTGGGCGCTTTCCACG 3240  
QY 3241 ACCTTGAGGGGAAGCCATGATCTGAATATAGTATATCTTGTAAATACGTGAACAA 3300  
DB ACCTTGAGGGGAAGCCATGATCTGAATATAGTATATCTTGTAAATACGTGAACAA 3300  
QY 3301 ACCAAACCGTTTGTGTAAGGGAAGCTAAATGATTTTAAAAATCTATGTTTAA 3360  
DB ACCAAACCGTTTGTGTAAGGGAAGCTAAATGATTTTAAAAATCTATGTTTAA 3360  
QY 3361 AATACTATGTAATCTTTTTCATCTATTAGTATATATTTATGATGGAATAAACTTTC 3420  
DB AATACTATGTAATCTTTTTCATCTATTAGTATATATTTATGATGGAATAAACTTTC 3420  
QY 3421 TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453  
DB TACTGTAAAAAATAAAAAAAAAAAAAAAAAAAAAA 3453

RESULT 2

US-09-872-136-1  
; Sequence 1, Application US/09872136  
; Patent No. US20020119545A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/872,136  
; FILING DATE: 01-Jun-2001

CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/208,786  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/09/021,324  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/07/977,451  
FILING DATE: 1992-11-19  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3453 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 112..3006  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 31..111  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 31..3009  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-872-136-1  
Query Match 100.0%; Score 3453; DB 10; Length 3453;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3453; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCGGCTGCTACCGCGCGCTCGGAGGCATCGCGGCTTGCGCGAGCGCAGCGACCG 60  
DB 1 GCGGCTGCTACCGCGCGCTCGGAGGCATCGCGGCTTGCGCGAGCGCAGCGACCG 60  
QY 61 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
DB 61 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
QY 121 CTGCTGCTGATCAAGTGTGTTTTTAATCAGTCATGAGAACAAATGGCTCATCAGCGGAAAG 180  
DB 121 CTGCTGCTGATCAAGTGTGTTTTTAATCAGTCATGAGAACAAATGGCTCATCAGCGGAAAG 180  
QY 181 CCATCATCGTACCGAATGTGGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGCGG 240  
DB 181 CCATCATCGTACCGAATGTGGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGCGG 240

241 CAGAGTGAAGGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC 300  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
241 CAGAGTGAAGGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATC 300  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
301 ACCTGCAAGTGCAGCTCGCCACCCAGGGAGCTTTCTCTGCTCTGGGTCTTTAAGCAC 360  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
301 ACCTGCAAGTGCAGCTCGCCACCCAGGGAGCTTTCTCTGCTCTGGGTCTTTAAGCAC 360  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
361 AGCTCCCTGGGCTGCCAGCGGCACCTTTGATTACAAAACAGAGGAATCGTTTCCATGGCC 420  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
361 AGCTCCCTGGGCTGCCAGCGGCACCTTTGATTACAAAACAGAGGAATCGTTTCCATGGCC 420  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
421 ATCTTGAACGTGACAGAGACCCAGGAGAGATACTACTCCATATTCAGAGCGAAGC 480  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
421 ATCTTGAACGTGACAGAGACCCAGGAGAGATACTACTCCATATTCAGAGCGAAGC 480  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
481 GCCAATACACAGTACTGTTACAGTGAATGTGAAGATACACAGCTGTATGTGCTAAGG 540  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
481 GCCAATACACAGTACTGTTACAGTGAATGTGAAGATACACAGCTGTATGTGCTAAGG 540  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
541 AGACCTTACTTTAGGAAGATGGAACCCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
541 AGACCTTACTTTAGGAAGATGGAACCCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
601 CCGAGGCCACTGTGGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
601 CCGAGGCCACTGTGGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
661 GGCCTGCTGTTCTCAGAAAGGAGAAAGTACTTTCATGAGTTGTTTCGGAACAGACATC 720  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
661 GGCCTGCTGTTCTCAGAAAGGAGAAAGTACTTTCATGAGTTGTTTCGGAACAGACATC 720  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
721 AGATGCTGCTAGAAATGCATGGCGCGAATGCACCAAGCTGTTTACCATAGATCTA 780  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
721 AGATGCTGCTAGAAATGCATGGCGCGAATGCACCAAGCTGTTTACCATAGATCTA 780  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
781 AACAGGCTCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGGAACCCCTGTGG 840  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
781 AACAGGCTCTCAGAGCACACTGCCCCAGTTATTCCTGAAAGTGGGGGAACCCCTGTGG 840  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
841 ATCAGGTGAAGCCCATCCATGTGAACCATGGATTTCGGGCTCAGCTGGGAGCTGGAAGAC 900  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
841 ATCAGGTGAAGCCCATCCATGTGAACCATGGATTTCGGGCTCAGCTGGGAGCTGGAAGAC 900  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
901 AAAGCCCTGGAGGAGGCGACCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 960  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
901 AAAGCCCTGGAGGAGGCGACCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATG 960  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
961 ATTCGGATTCTCTTGCCCTTTGTCTCTCCGTGGGAAGAACGACACCGGATATTACACC 1020  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
961 ATTCGGATTCTCTTGCCCTTTGTCTCTCCGTGGGAAGAACGACACCGGATATTACACC 1020  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1021 TGCTCTCTCAAGCACCCCGAGCTGAGCGTTGGTGGACCTCCCTAGAAAAAGGTTT 1080  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1021 TGCTCTCTCAAGCACCCCGAGCTGAGCGTTGGTGGACCTCCCTAGAAAAAGGTTT 1080  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1081 ATAAACGCTACCAAGTCCGCAAGAGAGTAAATGAAATGACCCGTCACGAAAAGTTCTGCTTC 1140  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1081 ATAAACGCTACCAAGTCCGCAAGAGAGTAAATGAAATGACCCGTCACGAAAAGTTCTGCTTC 1140  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1141 TCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGACGAGTCTTCTCTCAAGACCTCA 1200  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1141 TCAGTCAGGTTTAAAGCGTACCCACGAATCCGATGACGAGTCTTCTCTCAAGACCTCA 1200  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1201 TTTCTCTGTAACAGAGAGCGCTGGAGATGGGTACAGCATATCTAAATTTTGGCGATCAT 1260  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1201 TTTCTCTGTAACAGAGAGCGCTGGAGATGGGTACAGCATATCTAAATTTTGGCGATCAT 1260  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1261 AAGAAACAGCCAGAGAGTACATATCTATGCAGAAAATGATGACGCCCGAGTTCCACAAA 1320  
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1261 AAGAAACAGCCAGAGAGTACATATCTATGCAGAAAATGATGACGCCCGAGTTCCACAAA 1320  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1321 ATGTTACGCTGAATATAGAAAACCTCAAGTGTAGCAATGCTCTAGCCAGCCAG 1380



Sequence 3, Application US/09919408  
Patent No. US20020072077A1  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ImClone Systems Incorporated  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/919,408  
FILING DATE: 31-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/977,451  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEW-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3501 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 58..3039  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 139..3036  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 58..138  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-919-408-3  
Query Match 64.5%; Score 2227.2; DB 10; Length 3501;  
Best Local Similarity 79.6%; Pred. No. 0;

Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

QY	8	GGCTACCGCGCTCGGAGGCCATCGCGCGTTGGCGCAGCGAGCCGCGCGTGC	67
DB	35	GGGACCCCGCGCTCGGAGGCCATCGCGCGTTGGC---GCGGACGCGGACCGTGC	91
QY	68	TGCTGCTGTTGTTTTCAGTAATGATCTTGAGACCGTTACAAACAGACCTGCCTG	127
DB	92	CGCTGCTGCTGTTTTCGCAATGATATTTGGGACTATTACAAATCAAGATCGCTG	151
QY	128	TGATCAAGTGTGTTTAAATCAGTCATCAGAACAAATGCTCATCAGCGGGAAGCCAT	187
DB	152	TGATCAAGTGTGTTTAAATCAGTCATCAGAACAAATGCTCATCAGTGGGAAGTCAT	211
QY	188	CGTACCGAATGTCGCGAGGATCCCGAAGACCTCCAGTGTACCCCGAGGCGCAGAGT	247
DB	212	CATATCCCATGGTATCAGAAATCCCGAAGACCTCGGCTGCGTTCAGACCCGAGAGT	271
QY	248	AAGGACGGTATATGAAGCGGCCACCGTGGAGTGGCGGAGTCTGGGTCCATCACCTGC	307
DB	272	CAGGACAGTGTACGAAGCTGCCGCTGTGGAAGTGGATGATCTGCTTCCATCACACTGC	331
QY	308	AAGTGCAGTCCGCCACCCCGAGGGACCTTTCCCTCTGGGTCTTTAAGCACAGCTCCC	367
DB	332	AAGTGTGTCGATGCCCGAGGAAATTTCTGCTCTGGGTCTTTAAGCACAGCTCCC	391
QY	368	TGGGCTGCCAGCGCACTTTGATTTACAAACAGAGAAATCGTTTCCATGCGCCATCTGA	427
DB	392	TGAATGCCAGCCACATTTTGAATTTACAAACAGAGAGTGTGTTTCCATGCTATTTGA	451
QY	428	ACGTGACAGAGACCCAGGAGGAGATACCTACTTCCATATTCAGAGGGAAGCGCCAACT	487
DB	452	AAATGACAGAAACCCAAAGCTGGAGATACCTATTTTATTCAGAGTGAAGCTACCAAT	511
QY	488	ACACAGTACTGTTTCACAGTGAATGTAAGAGATACACAGCTGTATGCTTAGGAGACCTT	547
DB	512	ACAAATATTTGTTTACAGTGAATGTAAGAAATACCTGCTTACACATTAAGAGACCTT	571
QY	548	ACTTTAGGAAGTGGAAACACAGGATGCACCTGCTGTCATCTCCGAGGCTGTCGAGC	607
DB	572	ACTTTAGAAATGGAAACACAGGACCCCTGCTGTCATGTCATGAGTGTTCGACAGCTT	631
QY	608	CCACTGTGAGTGGGTGCTCTGACGCTCCACAGGAAAGCTGTAAAGAAAGAGCCCTG	667
DB	632	CGATCGTGAATGGTCTTTGCCATTCACAGGGGAAAGCTGTAAAGAAAGAGTCCAG	691
QY	668	CTGTTGTCAGAAAGGAGGAAAGTACTTATGAGTGTTCGACAGACACATCAGATGCT	727
DB	692	CTGTTGTTAAAGGAGGAAAGTGTCTCATGAATTTTGGGACGACATAGGCTGCT	751
QY	728	GTGCTAGAAATGCACTGGCGCGAATGCACCAAGCTGTTCCACCATAGATCTAAACAGG	787
DB	752	GTGCCAGAAATGAACCTGGGAGGAAATGCACAGGCTGTTCCAAATAGATCTAAATCAA	811
QY	788	CTCCTCAGACACACTGCCCCAGTTATTCCTGAAAGTGGGGAACCCCTTGTGGATCAGGT	847
DB	812	CTCCTCAGACACACTGCCCAATTTATTTCTAAAGTAGGGGAACCCCTTATGGATTAAGT	871
QY	848	GTAAGGCCATCCATGTGAACCATGGATTCGGGCTCACCTGGAGCTGGAAGCAAGGCC	907
DB	872	GCAAAGCTGTTTCATGTGAACCATGGATTCGGGCTCACCTGGGAAATAGAAACAAAGC	931
QY	908	TGGAGAGGCGAGCTACTTTTCAGATGAGTACTCTCTCCACAAACAGGACCATGATTCGA	967
DB	932	TCGAGGAGGCGAAGTACTTTTCAGATGAGTACTCTCTCCACAAACAGGACCATGATTCGA	991
QY	968	TTCTCTTGGCCTTTGCTCTTCCCTGGGAAGGAGACACCGGATTAATACACCTCTCTT	1027
DB	992	TTCTGTTGCTTTTGTATCATCAGTGGCAAGAAACGACACCGGATTAATACACCTCTCTT	1051
QY	1028	CCTCAAGACACCCAGGCGAGTACGCTGTTGGTGTGACCATCTCTAGAAAAAGGTTTAAAGC	1087
DB	1052	CTTCAAGCATCCCGAGTCAATCAGCTTTGGTTTACCATGCTAGGAAAGGATTTATAAATG	1111

1088 CTACGAGTCCGCAAGAGATGATGAAATGACCGGTACGAAAGTTCTGCTCTCAGTCA 1147  
1112 CTACCAATTCAGTGAAGATGATGAAATGACCAATGATGAAGATTTGTTTTCTGTGCA 1171  
1148 GGTTFAAAGCGTACCACGAATCCGATGCACGTGGATCTTCTCAAGCCTCATTTCCCT 1207  
1172 GGTTFAAAGCGTACCACGAATCCGATGCACGTGGATCTTCTCGAAATCATTTCCCT 1231  
1208 GTGACAGAGAGCGCTGGAGGATGGTACAGATATCTAAATTTGCGATCATAGAACA 1267  
1232 GTGACAAAGAGGCTTGTATAACGATACAGATATCCAAAGTTTGTCAATCATAGACCC 1291  
1268 AGCCAGGAGATPACATATCTATGACAGAAATGATGACCCCGAGTTTACCAAAATGTTCA 1327  
1292 AGCCAGGAGATATATATCTCATGACAGAAATGATGATGATGATGATGATGATGATGAT 1351  
1328 CGCTGAATATAAGAAAGAACCTCAAGTGTAGCAAAATGCTTCAGCCAGCGGTCTC 1387  
1352 CGCTGAATATAAGAAAGAACCTCAAGTGTAGCAAAATGCTTCAGCAAGTCAAGCGTCT 1411  
1388 GTTCTCTGTAGGCTACCCGCTACCCCTCTTGACCTGGAGAGTGTTCGGACAAATCTC 1447  
1412 GTTCTCGATGGATACCCATTTACCATCTTGACCTGGAGAGTGTTCAGCAAGTCTC 1471  
1448 CCAATTCACGAGGAGAAATCCAGAGGAGTGTGGAATAAAAGGCTAACAGAAAGTGT 1507  
1472 CCACTGCACAGAGATATCAGAGAGTGTGGAATAGAAGGCTAACAGAAAGTGT 1531  
1508 TTGGCCAGTGGTGTGAGCAGTACTCTAAATPATGAGTGAGCGCGGAAAGGCTTCTGG 1567  
1532 TTGGACAGTGGTGTGAGCAGTACTCTAAACATGAGTGAAGCCATAAAAGGTTCTCTGG 1591  
1568 TCAATGCTGTGCTGACATCTATGCGCAGCTCTTGGAACACCTCTTTTAAACTAC 1627  
1592 TCAAGTGTGTGACATACAAATTCCTTGGCAGATCTTGTGAGCAGTCTCTTTTAAACTCTC 1651  
1628 CAGGCGCTCTCCCTTTTCAAGACAACTCTCTCTATGCGACCTATGCGCTCTGTC 1687  
1652 CAGGCGCTCTCCCTTTTCAAGACAACTCTCTCTATGCGACCTATGCGCTCTGTC 1711  
1688 TCCCTTCATTTGTTCTCATTTGTTGATGCTGCACAAATACAAAAGCAATTTAGT 1747  
1712 TCCCTTCATTTGTTCTCATTTGTTGATGCTGCACAAATACAAAAGCAATTTAGT 1771  
1748 ACGAGATGACGTGCAGATGATCCAGGTGACTGGCGCCCTGGATACGAGTACTTCTAG 1807  
1772 ATGAAAGCAGCTACAGATGTTACAGGTGACCGGCTCTCAGATATGAGTACTTCTAG 1831  
1808 TTGACTTCAGGAGTATGAATGACCTTAAGTGGAGTTCCTCGAGAGAACTTAGAT 1867  
1832 TTGATTCAGAGAAATGAATATGATCAAAATGGAGTTTCCAGAGAAATTTAGAT 1891  
1868 TTGGAAGGCTCTGGGCTCTGGGCTTTCGGAGGCTGATGAACGCCAGCGCTATGGCA 1927  
1892 TTGGAAGGCTACTAGATCAGGTGCTTTGGAAGATGATGAACGCCAGCTTATGNA 1951  
1928 TTGTAAGGCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAAGCTGACA 1987  
1952 TTGTAAGGCGGAGTCTCAATTCAGGTGGCGGTGAAGATGCTAAAGAGAAAGCTGACA 2011  
1988 GCTGTGAAAAGAACTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047  
2012 GCTGTGAAAAGAGAGCTCATGTCAGAACTCAAGATGATGATGATGATGATGATGATGAT 2071  
2048 ACAACATCGTGAATCTGTGGGGGATGACACTGTGAGGGCGAGTACTGATTTTGT 2107  
2072 AGAATATTGTGAACCTGTGGGGGCTGCACACTGTGAGGACCAATTTACTTGATTTTG 2131  
2108 AATATTGTTGATGGTACCTCTCAACTACTTAAGAGTAAAAGAGAAAGTTCACA 2167  
2132 AATACTGTTGATGATGATCTCTCAACTATCTAAGAGTAAAAGAGAAATTTTACA 2191

2168 GGACATGGACAGAGATTTTAAAGAAACATAAATTCAGTTCTTACCTACTTTCCAGCAC 2227  
2192 GACATGGACAGAGATTTTCAAGAAACACAAATTCAGTTTACCCACATTTCCAAATCAC 2251  
2228 ATTCAAATTCAGATGCTGTTTACAGAAAGTTCAGTTTACACCCGCCCTTGGATCAGC 2287  
2252 ATCCAAATTCAGATGCTGTTTCAAGAAAGTTCAGATACACCCGCACTCGGATCAAA 2311  
2288 TCTCAGGTTTCAATGGATTCATCTCAAGATGATTTGAATATGAATGAATGAATGAATGA 2347  
2312 TCTCAGGTTTCAATGGATTCATCTCAAGATGATTTGAATATGAATGAATGAATGAATGA 2371  
2348 AGAGGTTGTCAGAAAGAGGAGGATTTGAACGTTGCTGAGCTTTTGAAGACCTCTCTT 2407  
2372 AAAGGCTG-----GAAGAGAGAGGACTTGAATGCTTACATTTGAAGATCTCTT 2425  
2408 GCTTTGCTTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAGTCTGTGTCTCACA 2467  
2426 GCTTTGCTTATCAAGTGGCCAAAGGATGGAATTTCTGGAATTTAAGTCTGTGTCTCACA 2485  
2468 GAGACTGTCAGCAGGAAATGCTTGTCTACCCACCGGAGGTTGTTGAAGATCTGTGACT 2527  
2486 GAGACTGTCAGCAGGAAATGCTTGTCTACCCACCGGAGGTTGTTGAAGATCTGTGACT 2545  
2528 TTGGACTGTCGCGAGACATCTGAGCAGCTTCCAGCTACGTCAGGGGCAACGACGCG 2587  
2546 TTGGATTTGCTCGAGATATCATGAGTATTCACACTATGTTGTCAGGGGCAATGCCGCTC 2605  
2588 TCCGCTGTAAGTGGATGGCAGCAGAGCTTATTTGAAGGATCTACACAAATCAAGAGT 2647  
2606 TGCTGTAAATGGATGGCCCGGAAAGCTGTTGAAGGATCTACACAAATCAAGAGT 2665  
2648 AGCTGCTGCTTACGACATCTCTCTGAGATATTTTACCTGGGTGTGAACCTTTACC 2707  
2666 ATGCTGCTGATGAAATTTACTGTGGAATCTTCTCAGTTGGTGTGAATCTTTACC 2725  
2708 CTGGCATCTCTGTCGACGCTAACTTCTATAAAGTATTCAGAGTGGATTTAAATGAGC 2767  
2726 CTGGCATCTCTGTCGATGCTAACTTCTACAACTGATTCAAATGATTTAAATGATC 2785  
2768 AGCCATTTCTATCCACAGAGGATATCTTTGTAATGCAATCTCTGCTGGCTTTTGAAT 2827  
2786 AGCCATTTCTATGTCACAGAGAAATATACATTAATGCAATCTCTGCTGGCTTTTGAAT 2845  
2828 CAGGAGGCGGCTATCTTCCCACTGACTTCAATTTTAGATGTCAGCTGGCAGG 2887  
2846 CAGGAGGCGGCTATCTTCCCTAAATTTGACTTCTGTTTGGATGTCAGCTGGCAGG 2905  
2888 CAGAAGAG-----CATGTATCAGAAACATCCATCATCTACC 2924  
2906 CAGAAGAGGATGATCAGATGTCAGTGGCTGTTTCCGAATGTCCTCAGACCTTACC 2955  
2925 AAAACAGGCGGCTTCCAGCAGAGAGCGGCTCAGAGC--CCAGTCCGCCACAGCGCCAGG 2983  
2966 AAAACAGGCGCTTCCAGCAGAGAGATGATTTGGGCTACTCTCTCCCGCAGGCTCAGG 3025  
2984 T-GAGATTCACAGAGAAAGATTTAGGAGGAGGCTTGGACCCGCC-----ACCCTAG 3038  
3026 TCGAAGATTCGTAAGAGAAACATTTAGTTTAAAGGACTTCATCCCTCCACTATCCCTAA 3085  
3039 CAGGCTGTAGACGCGCAGAGATAGCTTCCCTCTCTCT--GAGGAGCGCCCTACAGCG 3096  
3086 CAGGCTGTAGATTTACCAAAACAGATTAATTTCTACATAAAGAAATCTATTATCAAC 3145  
3097 CAGTGTCTCGCTGGCTTTTCTAGATGCTGTCTGCTTACTTCT-----CAAGTGA 3149  
3146 TCTGCTTCCAGCAGCTTTCTCAGAGCGCTCTGCTTACTCTTGTGTTTCAAGGGA 3205  
3150 CTTCTATAAATCAAACTCTCTCTCAGCAGCGGAGGAGCCAAATTAATGAGCTTGTGG 3209  
3206 CTTTGTAAATCAAACTCTCTCTCAGAGGAGGAGGAGCTGATAATGAACCTTTTATG 3265  
3210 TGAGCCCGCTTACCTTGGGGCTTTTCCAGGAGCTTTGAGGGGAAAGCCATGTATCTGAAA 3269

Db 3266 GAGCATTGATCTGCATCCAAAGGCTTCTCAGCGGGCTTGAGTGAATTTGCTACCTGAAG 3325  
QY 3270 TATAGTATATCTGTGTAATACGTGAACAAACACCGTCTTTTGTCTAAGGGAAGC 3329  
Db 3326 TACAGTATATCTGTGTAATACATACAAACAAA-----AGCATTTTGTGAAGGAAGC 3378  
QY 3330 TAAATATGATTTTAAATACTATGTTTAAATACATATGTAACCTTTTTCATCTATTAG 3389  
Db 3379 TAATATGATTTT--AAGTCTATGTTTAAATAATATGTAATTTTTCAGCTATTAG 3436  
QY 3390 TGATATATTTATGGATGGAATAAATCTTCTACTGTAAATAAATAAATAAATAAATAA 3449  
Db 3437 TGATATATTTATGGTGGGAATAAATTTCTACTACAGAAATAAATAAATAAATAAATAA 3496  
QY 3450 AAAA 3453  
Db 3497 AAAA 3500

## RESULT 4

US-09-872-136-3

Sequence 3, Application US/09872136

Patent No. US20020119545A1

GENERAL INFORMATION:

APPLICANT: Lemischka, Ihor R.

TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESS/SEE: Imclone Systems Incorporated

STREET: 180 Varick Street

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10014

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION NUMBER: US/09/872,136

FILING DATE: 01-Jun-2001

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/208,786

FILING DATE: &lt;Unknown&gt;

APPLICATION NUMBER: US/09/021,324

FILING DATE: &lt;Unknown&gt;

APPLICATION NUMBER: US/07/977,451

FILING DATE: 1992-11-19

APPLICATION NUMBER: US 07/906,397

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: US PCT/US92/05401

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: TW 81102961

FILING DATE: 15-APR-1992

APPLICATION NUMBER: US PCT/US92/02750

FILING DATE: 02-APR-1992

APPLICATION NUMBER: US 07/813,593

FILING DATE: 24-DEC-1991

APPLICATION NUMBER: US 07/793,065

FILING DATE: 15-NOV-1991

APPLICATION NUMBER: US 07/728,913

FILING DATE: 28-JUN-1991

APPLICATION NUMBER: US 07/679,666

FILING DATE: 02-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Feit, Irving N.

REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: LEM-3-7P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 3501 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

FEATURE:

NAME/KEY: CDS

LOCATION: 58..138

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: 139..3036

FEATURE:

NAME/KEY: sig\_peptide

LOCATION: 58..138

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-872-136-3

Query Match 64.5%; Score 2227.2; DB 10; Length 3501;

Best Local Similarity 79.6%; Pred. No. 0;

Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

QY 8 GGCTACCGCGGCTCCGGAGGCGCATCGGGCGCTTGGCGAGCGAGCGAGCGCGGCTGC 67  
Db 35 GGGGACCGCGGCTCCGGAGGCGCATCGGGCGCTTGGC---GCGGACGCGGCGACCGTGC 91  
QY 68 TGCTGCTTGTGTTTGTTCAGTAATGATTTTGTAGACCGCTTACAAACCAAGACCTGCCG 127  
Db 92 CGCTGCTGCTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCTG 151  
QY 128 TGATCAAGTGTGTTTATATCAGTCATGAGAAATGGCTCATCAGCGGGAAGCCATCAT 187  
Db 152 TGATCAAGTGTGTTTATCAATCATGAGAAATGATTTCATCAGTGGGGAAGTCAATCAT 211  
QY 188 CGTACCGAATGCTGCGAGGATCCCGAGAGACCTCCAGTGTACCCGAGCGCGCAGAGTG 247  
Db 212 CATATCCCATGATATCAGAAATCCCGAGAGACCTCGGTGTGCGTTGAGACCCAGAGCT 271  
QY 248 AAGGACCGGTATATGAAGCGGCACCGTGGAGTGGCGGAGCTGTGGTTCATCACCGTGC 307  
Db 272 CAGGACACTGTACGAAGCTGCGCTGTGGAAGTGTATCTGCTTCATCACACTGC 331  
QY 308 AAGTGCAGTCCGCCACCGCGGAGCTTTCCTGCTCTGGGTCTTTAAGCAGAGCTCCC 367  
Db 332 AAGTGCCTGCTGATGCCCGAGGGAACATTTCTGCTCTGCGGTCTTTAAGCAGAGCTCCC 391  
QY 368 TGGGCTGCCAGCGCGCACTTTGATTTACAAAACAGAGGAATCGTTTCATGCCCATCTTGA 427  
Db 392 TGAATTCGCCAGCCACATTTGATTTACAAACAGAGAGAGTGTGTTCCATGTCATTTTGA 451  
QY 428 AGCTGACAGAGACCCAGGAGGAATACCTACTCCATATTCAGAGGACGCGGCGCAACT 487  
Db 452 AAATGACAGAAACCCCAAGCTGGAGAATACCTACTTTTATTCAGAGTGAAGCTACCAAT 511  
QY 488 ACACAGTACTGTTTCAGTGAATGTAAGAGATACAGCTGTATGTCTTAAAGAGACCTT 547  
Db 512 ACACAATATTTTACAGTGAGTATAAGAAATACCTGCTTTACACATTAAGAGACCTT 571  
QY 548 ACTTTAGGAAGATGGAAACCCAGAGTCCACTGCTGCTCATCTCCGAGGGTGTTCGGGAGC 607  
Db 572 ACTTTAGAAATGGAAACCCAGAGCGCCCTGCTGCTCATATCTGAGAGCGTTCCAGAGC 631  
QY 608 CCACTGCTGAGTGGTGTCTGTCAGCTCCCAAGGGAAGCTGTAAAGAGAGAGCGGCTG 667  
Db 632 CGATCGTGAATGGTGTCTTTGGGATTCACAGGGGGAAGAGCTGTAAAGAGAGAGTCCAG 691  
QY 668 CTGTTGTCAGAAAGAGGAAAGGTAATCTTCAATGAGTTGTTTCGGAACACATCATGCT 727





2888 CAGAAGAAG-----CATGTATCAGAACATCATCATCACC 2924  
|||||  
2906 CAGAAGAAGCATGTATCAGATGTGGATGGCGGTGTTTCGGAATGCTCTCACCTACC 2965  
|||||  
2925 AAACAGGGCGCCCTCAGCAGAGAGGGGCTCAGAGC-CCAGTCGCCACAGCCGAGG 2983  
|||||  
2966 AAACAGGGGACCTTTCAGCAGAGAGATGGATTTGGGGCTACTCTCTCGCGAGGCTCAGG 3025  
|||||  
2984 T-GAAGATTCACAGAGAAGAGTTAGCGAGAGGCGCTTGGACCCCGCC-ACCTAG 3038  
|||||  
3026 TCGAAGATTCGTAAGAGCAATTTAGTTTAAAGGACTTCATCCCTCCACCTATCCCTAA 3085  
|||||  
3039 CAGGCTGTAGACCGCAGAGCCAGATAGCTCGCTCT--GAGGAAGCGCCCTCAGCG 3096  
|||||  
3086 CAGGCTGTAGATACCAACAGAGTAATTTTCATCAGCTAAAGAAATCTATTATCAAC 3145  
|||||  
3097 CTTGCTGCTGGGACTTTCTCTAGATGCTGTGCTGCTTACTC-----CAAAGTGA 3149  
|||||  
3146 TCGTCTTCCACGAGACTTTCTCTAGAGCGCTGCTGTTTACTCTTTGTTTCAAAGGGA 3205  
|||||  
3150 CTTCTATAAATCAACCTCTCTCGCAGAGGGGAGAGCCAAATAGAGACTTGTGG 3209  
|||||  
3206 CTTTGTAAATCAATCATCTGTCTACAAGGCGAGGAGCTGATTAATGAATTTATTG 3265  
|||||  
3210 TGAGCCGCGCTACCTGGGGCGCTTTCACAGAGCTTGAGGGGAAAGCCATGTATCTGAAA 3269  
|||||  
3266 GAGCATGTATCTGATCCAGGCTTCTCAGCGCGCTTGAGUAGATTTGTACCTGAG 3325  
|||||  
3270 TATAGTATATCTTGTAAATAGCTGAACAAACCAACCCGTTTGTGCTAAGGGAAGC 3329  
|||||  
3326 TACAGTATATCTTGTAAATACATAAAACAA-AGCATTTTGTGCTAAGGGAAGC 3378  
|||||  
3330 TAAATATGATTTTAAATATCTATGTTTAAATATCTATGTAACCTTTTCTATCTATTAG 3389  
|||||  
3379 TAATATGATTTT--AAGTCTATGTTTAAATATATGTAATTTTTCAGCTATTAG 3436  
|||||  
3390 TCATATATTTTATGATGAATAAATCTTCTACTTAAATATAAATATAAATATAAATATAA 3449  
|||||  
3437 TCATATATTTTATGTTGGTGGTAATAAATTTCTACTACAGAAATATAAATATAAATATAA 3496  
|||||  
3450 AAAAA 3453  
|||||  
3497 AAAAA 3500  
|||||

RESULT 5

US-09-982-610-22  
; Sequence 22, Application US/09982610  
; Patent No. US2002014620A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.  
Bennett, Brian D.  
Goeddel, David  
Lee, James M.  
Matthews, William  
Tsai, Siao Ping  
Wood, William I.

TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/982,610  
FILING DATE: 17-Oct-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/446,648  
FILING DATE: 1996-MAY-23  
APPLICATION NUMBER: 08/22616  
FILING DATE: 04-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 40,378  
REFERENCE/DOCKET NUMBER: P0821P3PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/952-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3120 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 22:  
US-09-982-610-22

Query Match 60.3%; Score 2081.8; DB 10; Length 3120;  
Best Local Similarity 80.9%; Pred. No. 0;  
Matches 2513; Conservative 0; Mismatches 557; Indels 37; Gaps 6;

QY 48 GCGCAGCGACCGCGCGCTGCTGCTGTTGTTTTCAGTATGATTTCTTGAGACCGT 107  
Db 15 GCGCAGCGCGCGCGCTGCTGCTGTTGTTTTCAGTATGATTTCTTGAGACCGT 74  
QY 108 TACAAACCAAGAGCTGCTGCTGATCAAGTCGTTTAAATCAGTCATCAGACATGCGTC 167  
Db 75 TACAAATCAAGATCTGCTGCTGATCAAGTCGTTTAAATCAGTCATCAGACATGATTC 134  
QY 168 ATCAGCGGGAAGAACCCATCATCTGATCCGAATGGTGGCAGATCCCGAAGACCTCCAGTG 227  
Db 135 ATCAGTGGGAAGTATCATCATATCCATGGTATCAGATCCCGGGAAGACCTCGGCGT 194  
QY 228 TACCCGAGCGCCAGAGTGAAGGCGGCTATATGAAGCGGCCACCGTGGAGGTGGCGGA 287  
Db 195 TCGGTTGAGACCCAGAGCTCAGGCGAGCTGACGAAGCTGCCGCTGTGGAAGTGGATGT 254  
QY 288 GTCTGGGTCCATCACCCTGCAAGTGCAGCTGCCACCCAGGGGACCTTTCCTGCTCTG 347  
Db 255 ATCTGCTTCATCAGCTGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 314  
QY 348 GGTCTTTAAGCACAGCTCCCTGGGCTGCCAGCCGACCTTTGATTTACAAACACAGAGGAA 407  
Db 315 GGTCTTTAAGCACAGCTCCCTGGTATGCCAGCCACATTTTATTTACAAACACAGAGGAGT 374  
QY 408 CGTTTCCATGGCCATCTTGAACGTGACAGAGACCCAGGAGGAGAAATACCTACTCATAT 467  
Db 375 TGTTCATGCTCATTTTGAATAATGACAGAAACCCAAAGCTGGAGATACCTACTTTTAT 434  
QY 468 TCAGAGGACGCGGCCCACTACACAGTACTGTTTACAGTGAATGATGATGATGATGATGAT 527  
Db 435 TCAGAGTGAAGCTACCAATTTACCAATTTTACAGTGAATGATGATGATGATGATGATGAT 494  
QY 528 GTATGTCTAAGGAGACCTTACTTTAGGAAGATGGAAGAACCCAGGATGCTGCTGCTGCT 587  
Db 495 TTACACATTAAGAGAGCTTACTTTAGAAATGGAAGAACCCAGGACCCCTGGTCTGCAT 554  
QY 588 CTCGAGGGTGTTCGGAGGCCACTGTGTGGTGGTGGTGTCTCTGAGCTCCACAGGGAAG 647  
Db 555 ATCTGAGAGCGTCCAGAGCGGATCTCTGGAATGGTGTCTTGGCATTACAGAGGGGAAG 614  
QY 648 CTGTAAGAAGAGGCGCTGCTGTTGTCAAGAGGAGGAAAGTACTTCATGATGTTGT 707  
Db 615 CTGTAAGAAGAGAGTCCAGCTGTTGTTAAAGAGGAGGAAAGTGTCTTCAATGATTTAT 674



```
Db 2829 AGGATGTGAGCTGGCAGATGCAGAAAGCGATGTATCAGAATGTGGATGCCGCTGTTTC 2888
QY 2905 AGACATCCATCCATCTACCAAAACAGCGCGCCCTCAGCAGAGCGCGGCTCAGAGC- 2963
Db 2889 GGAATGTCTCAGACCTACCAAAACAGCGCGCCCTTTCAGCAGAGATGSGATTTGGGCT 2948
QY 2964 CCAGTCGCGCAGCGCCAGGT- GAAGATTCACAGAGAAAGTTAGCGAGGCGCTTG 3022
Db 2949 ACTCTCCGAGCTCAGGTCGAGATTCGAGAGGACAAATTTAGTTTAAAGGACTTC 3008
QY 3023 GACCC- ---CCGACCCCTAGCAGGCTGTAGCCGCGAGCAGCAAGATAGCCTCGCCTCT- 3077
Db 3009 ATCCCTCAGCTATCCCTCAGCAGGCTGTAGATACCAAAACAGGTTAATTTCACTACTA 3068
QY 3078 -GAGGAAGCGCCCTACAGCGGCTGCTTCGCTGCGTTCGACTTTCTCTAGA 3123
Db 3069 AAGAAAATCTATTATCAACTGCTCTCCAGCAGACTTTTCTCTAGA 3115

RESULT 6
US-09-796-692-7538
; Sequence 7538, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7538
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-7538

Query Match 13.5%; Score 466.6; DB 9; Length 645;
Best Local Similarity 84.5%; Pred. No. 8.7e-118;
Matches 550; Conservative 0; Mismatches 94; Indels 7; Gaps 2;

QY 2035 CTGGACACCATGACAAACATCGTGAATCTGCTGGGGCATGCACACTGTCTAGGGCCAGTG 2094
Db 1 CTGGGAAGCCACGAGAAATATTGTGAACCTGCTGGGGCGTGCGACACTGTCTAGGACCAATT 60
QY 2095 TACTTGATTTTGAATATTGTGCTATGCTGACCTCTCACTACCTAGAGAGTAAAGA 2154
```

```
Db 61 TACTTGATTTTGAATATTGTGCTATGCTGATCTCTCACTATCTAAGAGTAAAGA 120
QY 2155 GAGAAGTTTTCAGAGACATGGACAGAGATTTTAAAGAACATAATTTTCAGTTCTTACCT 2214
Db 121 GAAAATTTTCAGAGGACTTGGACAGAGATTTTCAAGAACACAAATTTTCAGTTTTCACCC 180
QY 2215 ACTTTCCAGGACATTTCAAATTCAGCATGCTGTTTACAGAGAACTTCAGTTTACACCCG 2274
Db 181 ACTTTCCAAATCAGATCCAAATTTCCAGCATGCTGTTTCAAGAGAACTTCAGATACACCCG 240
QY 2275 CCCTTGGTACGCTCTCAGGTTCAATGGGAATTCATTTCTCAAGATGAGATGAA 2334
Db 241 GACTCGGATCAATCTCAGGCTTCTAGGGAATTCATTTCACTCTCAAGATGAAATGAA 300
QY 2335 TATGAAAACAGAGAGGCTGCGCAGAGAGAGGAGAGATTTTGAACGCTGCTGACGTTT 2394
Db 301 TATGAAAACCAAAAGGCTG-----GAAGAGAGGAGGACTTGAATGTGCTTACATTT 354
QY 2395 GAAGACCTCTCTTGTGCTTTCGCTACCAAGTGCGCAAGGATGGAATTCCTGGAGTTCAAG 2454
Db 355 GAAGATCTTCTTGTGCTTTCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAAATTAAG 414
QY 2455 TCGTGTGTCCACAGAGACCTGCGCAGCAGGAATGTGTTGCTCACCACCGGAAGTGGTG 2514
Db 415 TCGTGTGTTCACAGAGACCTGCGCCGCGCAGGAGCTGTTGTACCCACCGGAAGTGGTG 474
QY 2515 AAGATCTGTGACTTTTGGACTGCGCGCAGACATCTCTGAGGAGCTCAGCTGCTGACG 2574
Db 475 AAGATATGTGACTTTTGGATTTGCTCGAGATATCATGAGTATGATTCACCACTATGTTG 534
QY 2575 GCGACGCGAGGCTGCGGCTGAGTGGATGGCGGAGACATCTCTGAGGAGCTCAGCTGCTG 2633
Db 535 GCGAATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594
QY 2634 CACAATCAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2684
Db 595 CACCATTAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645
```

## RESULT 7.

```
US-10-040-862-7538
; Sequence 7538, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
; FILE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
```

;; PRIORITY NUMBER: US 60/222,903  
;; PRIORITY FILING DATE: 2000-08-03  
;; PRIORITY APPLICATION NUMBER: US 60/223,416  
;; PRIORITY FILING DATE: 2000-08-04  
;; PRIORITY APPLICATION NUMBER: US 60/223,378  
;; PRIORITY FILING DATE: 2000-08-07  
;; PRIORITY APPLICATION NUMBER: US 09/796,692  
;; PRIORITY FILING DATE: 2001-03-01  
;; NUMBER OF SEQ ID NOS: 10467  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 7538  
;; LENGTH: 645  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-040-862-7538

Query Match 13.5%; Score 466.6; DB 9; Length 645;  
Best Local Similarity 84.5%; Pred. No. 8.7e-115;  
Matches 550; Conservative 0; Mismatches 94; Indels 7; Gaps 2;  
QY 2035 CTGGACACCATGACAACTGCTGAATCTGCTGGGGGATGCACACTGTCAGGGCCAGTG 2094  
DB 1 CTGGGACCCACGAGATATTTGTAACCTGCTGGGGGATGCACACTGTCAGGACCAATT 60  
QY 2095 TACTTGATTTTGAATATTTGCTATGCTGACCTCTCAACTACCTTAAGAGTAAAGA 2154  
DB 61 TACTTGATTTTGAATATTTGCTATGCTGATCTCTCAACTATCTTAAGAGTAAAGA 120  
QY 2155 GAGAGTTTCACAGGACATGACAGAGATTTTAAAGAACATATTTTCACTTACCTT 2214  
DB 121 GAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTTACCC 180  
QY 2215 ACTTTCCAGGACATCAAAATCCAGCATGCTGCTTCCAGGAGGATTTTACACCCG 2274  
DB 181 ACTTTCCAGTCAATCAATCTCAGGCTTCTAGGAAATTCATTTCACTCTGAAGATGA 240  
QY 2275 CCCTTGATCAGCTCTCAGGCTTCAATGGGAATCAATTTCTGAAGATGAATGAA 2334  
DB 241 GACTCGGATCAAAATCTCAGGCTTCTAGGAAATTTCTGAAGATGAATGAA 300  
QY 2335 TATGAAACCCAGAGAGCTGCGGACAGAGATTTTAAAGGAGGATTTTCACTTAC 2394  
DB 301 TATGAAACCCAGAGAGCTGCGGACAGAGATTTTAAAGGAGGATTTTCACTTAC 354  
QY 2395 GAAAGACCTCTTGTGCTTCCGTTACCAAGTGCGCAAGGATTTTCTCTGGAGATTT 2684  
DB 595 CACCATTAAGAGTATGCTGCTGATATGGAATATTTTCTGTTGGAATCTT 645

RESULT 8  
US-09-796-692-8085  
;; Sequence 8085, Application US/09796692  
;; Publication No. US20020198362A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Gaiger, Alexander  
;; APPLICANT: Algate, Paul A.  
;; APPLICANT: Mannion, Jane

;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER  
;; FILE OF INVENTION: HEMATOLOGICAL MALIGNANCIES  
;; FILE REFERENCE: 2077.001200  
;; CURRENT APPLICATION NUMBER: US/09/796,692  
;; CURRENT FILING DATE: 2001-03-01  
;; PRIORITY APPLICATION NUMBER: 60/186,126  
;; PRIORITY FILING DATE: 2000-03-01  
;; PRIORITY APPLICATION NUMBER: 60/190,479  
;; PRIORITY FILING DATE: 2000-03-17  
;; PRIORITY APPLICATION NUMBER: 60/200,545  
;; PRIORITY FILING DATE: 2000-04-27  
;; PRIORITY APPLICATION NUMBER: 60/200,303  
;; PRIORITY FILING DATE: 2000-04-28  
;; PRIORITY APPLICATION NUMBER: 60/200,779  
;; PRIORITY FILING DATE: 2000-04-28  
;; PRIORITY APPLICATION NUMBER: 60/200,999  
;; PRIORITY FILING DATE: 2000-05-01  
;; PRIORITY APPLICATION NUMBER: 60/202,084  
;; PRIORITY FILING DATE: 2000-05-04  
;; PRIORITY APPLICATION NUMBER: 60/206,201  
;; PRIORITY FILING DATE: 2000-05-22  
;; PRIORITY APPLICATION NUMBER: 60/218,950  
;; PRIORITY FILING DATE: 2000-07-14  
;; PRIORITY APPLICATION NUMBER: 60/222,903  
;; PRIORITY FILING DATE: 2000-08-03  
;; PRIORITY APPLICATION NUMBER: 60/223,416  
;; PRIORITY FILING DATE: 2000-08-04  
;; PRIORITY APPLICATION NUMBER: 60/223,378  
;; PRIORITY FILING DATE: 2000-08-07  
;; NUMBER OF SEQ ID NOS: 9597  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 8085  
;; LENGTH: 610  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-796-692-8085

Query Match 13.2%; Score 456.4; DB 9; Length 610;  
Best Local Similarity 85.1%; Pred. No. 5.3e-115;  
Matches 524; Conservative 0; Mismatches 86; Indels 6; Gaps 1;

QY 2035 CTGGACACCATGACAACTGCTGAATCTGCTGGGGGATGCACACTGTCAGGGCCAGTG 2094  
DB 1 CTGGGACCCACGAGATATTTGTAACCTGCTGGGGGATGCACACTGTCAGGACCAATT 60  
QY 2095 TACTTGATTTTGAATATTTGCTATGCTGACCTCTCAACTACCTTAAGAGTAAAGA 2154  
DB 61 TACTTGATTTTGAATATTTGCTATGCTGATCTCTCAACTATCTTAAGAGTAAAGA 120  
QY 2155 GAGAGTTTCACAGGACATGACAGAGATTTTAAAGAACATATTTTCACTTACCTT 2214  
DB 121 GAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTTACCC 180  
QY 2215 ACTTTCCAGGACATCAAAATCCAGCATGCTGCTTCCAGGAGGATTTTCACTTACACCCG 2274  
DB 181 ACTTTCCAGTCAATCAATCTCAGGCTTCTAGGAAATTCATTTCTGAAGATGAATGAA 2334  
QY 2275 CCCTTGATCAGCTCTCAGGCTTCAATGGGAATCAATTTCTGAAGATGAATGAA 2394  
DB 241 GACTCGGATCAAAATCTCAGGCTTCTAGGAAATTTCTGAAGATGAATGAA 300  
QY 2335 TATGAAACCCAGAGAGCTGCGGACAGAGATTTTAAAGGAGGATTTTCACTTAC 2394  
DB 301 TATGAAACCCAGAGAGCTGCGGACAGAGATTTTAAAGGAGGATTTTCACTTAC 354  
QY 2395 GAAAGACCTCTTGTGCTTCCGTTACCAAGTGCGCAAGGATTTTCTCTGGAGATTT 2454  
DB 355 GAGATCTCTTGTGCTTCCGTTACCAAGTGCGCAAGGATTTTCTCTGGAGATTT 414  
QY 2455 TCGTGTGTCCACAGAGACTTGGCAGCAGAGATTTTGGTTCACCCACGGAAGGTTG 2514  
DB 415 TCGTGTGTCCACAGAGACTTGGCAGCAGAGATTTTGGTTCACCCACGGAAGGTTG 474

```
QY 2515 AAGATCTGTACTTTGGACTGGCCGACAGACATCTGACGACTCCAGCTACGTCAGT 2574
||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 475 AAGATATGTACTTTGGATGGCTCGATATATCATGATGATTCACACTATGTTGTCAGG 534
||||| ||||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2575 GGCAACGACGCGCTGCGGCTGAAGTGGATGGACCCGAGAGCTTATTTGAAGGGATCTAC 2634
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 535 GGCAATGCCGCTGCTGCTGTAATGATGGATGGCCCGGAAAGCCTGTTTGAAGGCATCTAC 594
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2635 ACAATCAAGATGACG 2650
||| ||||||| |||
Db 595 ACCATTAAGAGTGATG 610
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
US-10-040-862-8085
; Sequence 8085, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8085
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-8085

Query Match 13.2%; Score 456.4; DB 9; Length 610;
Best Local Similarity 85.1%; Pred. No. 5.3e-115;
Matches 524; Conservative 0; Mismatches 86; Indels 6; Gaps 1;

QY 2035 CTGGGACACCATGACAACTCGTGAATCTGCTGGGGCATGCACTGTGACGGCCAGTG 2094
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 CTGGGAAGCCACGAGAAATATTGGAACCTGCTGGGGCGTGCACATGTGACGACCAATT 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2095 TACTTGATTTTGAATATTGTTGCTATGTGTGACTCTCTCAACTTACCTAAGTAAAGAA 2154
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TACTTGATTTTGAATATTGTTGCTATGTGTGACTCTCTCAACTTACCTAAGTAAAGAA 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
QY 2155 GAGAGTTTTCAGGACATGGACAGATTTTTTAAGGACATATATTTCAGTTCTTACCCT 2214
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 121 GAAATATTCACAGGACTTGGACAGATTTTCAAGGACACAAATTCAGTTTACCCTC 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 2215 ACTTTCAGGACATCAAAATTCAGCATGCTGTTTACGAGAGTTCAGTTACACCCG 2274
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 ACTTTCCAATCACATCAAAATTCAGCATGCTGTTTCAAGAGAAAGTTCAGATACACCCG 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 2275 CCCTTGGATCAGCTCTCAGGGTTCAATGGGAATTCAAATTCATTCTCAAGATGAGATTGAA 2334
||| ||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 241 GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTCTCAAGATGAAATTGAA 300
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 2335 TATGAAAACACGAGAGGCTGGCAGAAAGAGGAGGAAAGATTTTGAACGTGCTGACGTTT 2394
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 301 TATGAAAACCAAAAGGCTGGAAGAA-----GAGGAGGACTTGAATGTGCTTACATTT 354
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 2395 GAAGACCTCTTTTGGTGGTACCAAGTGGCCAAAGGATGGAATTCCTGGAGTTCAAG 2454
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 355 GAAGATCTTCTTTGCTTTCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTAAG 414
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 2455 TCGTGTGCCACAGAGACCTGGCAGCCAGGAATGTTGTTGTCACCCAGGAAGTGGTG 2514
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 415 TCGTGTGTTCAAGAGACCTGGCCGCGGAGGAGGATGCTTGTACCCAGGGAAGTGGTG 474
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 2515 AAGATCTGTGACTTTGGACTGGCCGACAGACATCTGAGCGACTTCCAGTACGTCAGG 2574
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 475 AAGATATGTGACTTTGGATGGCTCGATATATCATGATGATTCACACTATGTTGTCAGG 534
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 2575 GGCAACGACGCGCTGCGGCTGAAGTGGATGGCAGCCGAGAGCTTATTTGAAGGATCTAC 2634
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 535 GGCAATGCCGCTGCTGCTGTAATGATGGATGGCCCGGAAAGCCTGTTTGAAGGCATCTAC 594
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 2635 ACAATCAAGATGACG 2650
||| ||||||| |||
Db 595 ACCATTAAGAGTGATG 610
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
US-09-796-692-8927
; Sequence 8927, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
```

; PRIOR APPLICATION NUMBER: 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; NUMBER OF SEQ ID NOS: 9597  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8927  
; LENGTH: 605  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-796-692-8927

Query Match 13.2%; Score 454.6; DB 9; Length 605;  
Best Local Similarity 85.3%; Pred. No. 1.7e-114;  
Matches 521; Conservative 0; Mismatches 84; Indels 6; Gaps 1;  
QY 2035 CTGGGACACCATGCAATCTGCTGGGGCGTGCACACATGTCAGGCGCAGTG 2094  
Db 1 CTGGGAAGCCACGAGATATTTGAACCTCTGGGGCGTGCACACATGTCAGGACCAATT 60  
QY 2095 TACTTGATTTTGAATATTTGCTATGCTGATGCTCAACTCTCACTAAGAAAGTAAAGA 2154  
Db 61 TACTTGATTTTGAATATTTGCTATGCTGATGCTCAACTCTCACTAAGAAAGTAAAGA 120  
QY 2155 GAGAAGTTTTCACAGGACATGACAGAGATTTTAAAGAAATATTTCACTTACCTT 2214  
Db 121 GAAAAATTTTCACAGGACTTGGCAGAGATTTTAAAGAAACACAAATTTTCACTTAC 180  
QY 2215 ACTTTCAGGACATTCAAATTCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2274  
Db 181 ACTTTCAGGACATTCAAATTCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
QY 2275 CCCTTGATGAGTCTCAGGCTTCAATGGAATTCATCTCAAGATGAGATTGAA 2334  
Db 241 GACTCGGATCAATCTCAGGCTTCAATGGAATTCATCTCAAGATGAGATTGAA 300  
QY 2335 TATGAAACACGAGAGGCTGGCAGAGATTTTAAAGAAAGTAAAGA 2394  
Db 301 TATGAAACACGAGAGGCTGGCAGAGATTTTAAAGAAAGTAAAGA 354  
QY 2395 GAAGACTCTCTTCTGCTTTCAGGCTTCAAGTGGCAGAGATTTTCAAGTGGCAGAG 2454  
Db 355 GAAGATCTCTTCTGCTTTCAGGCTTCAAGTGGCAGAGATTTTCAAGTGGCAGAG 414  
QY 2455 TCCTTGATTTTGAATATTTGCTATGCTGATGCTCAACTCTCAAGTGGCAGAG 2514  
Db 415 TCCTTGATTTTGAATATTTGCTATGCTGATGCTCAACTCTCAAGTGGCAGAG 474  
QY 2515 AGATCTGCTGCTTTCAGGCTTCAAGTGGCAGAGATTTTCAAGTGGCAGAG 2574  
Db 475 AGATCTGCTGCTTTCAGGCTTCAAGTGGCAGAGATTTTCAAGTGGCAGAG 534  
QY 2575 GSCAACGCGGCTGCGGCTGAAGTGGATGCGACATCTGAGGCTTCAAGTGGCAGAG 2634  
Db 535 GSCAACGCGGCTGCGGCTGAAGTGGATGCGACATCTGAGGCTTCAAGTGGCAGAG 594  
QY 2635 ACAATCAAGAG 2645  
Db 595 ACATTAAGAG 605

RESULT 11  
US-10-040-862-8927  
; Sequence 8927, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; TITLE OF INVENTION: Hematological Malignancies  
; FILE REFERENCE: 014058-013520US  
; CURRENT APPLICATION NUMBER: US/10/040,862

; CURRENT FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: US 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: US 09/796,692  
; PRIOR FILING DATE: 2001-03-01  
; NUMBER OF SEQ ID NOS: 10467  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8927  
; LENGTH: 605  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-040-862-8927

Query Match 13.2%; Score 454.6; DB 9; Length 605;  
Best Local Similarity 85.3%; Pred. No. 1.7e-114;  
Matches 521; Conservative 0; Mismatches 84; Indels 6; Gaps 1;  
QY 2035 CTGGGACACCATGCAATCTGCTGGGGCGTGCACACATGTCAGGCGCAGTG 2094  
Db 1 CTGGGAAGCCACGAGATATTTGAACCTCTGGGGCGTGCACACATGTCAGGACCAATT 60  
QY 2095 TACTTGATTTTGAATATTTGCTATGCTGATGCTCAACTCTCAAGTGGCAGAG 2154  
Db 61 TACTTGATTTTGAATATTTGCTATGCTGATGCTCAACTCTCAAGTGGCAGAG 120  
QY 2155 GAGAAGTTTTCACAGGACATGACAGAGATTTTAAAGAAACATAATTTCACTTACCTT 2214  
Db 121 GAAAAATTTTCACAGGACTTGGCAGAGATTTTCAAGGAAACACAAATTTTCACTTAC 180  
QY 2215 ACTTTCAGGACATTCAAATTCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2274  
Db 181 ACTTTCAGGACATTCAAATTCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
QY 2275 CCCTTGATGAGTCTCAGGCTTCAATGGAATTCATCTCAAGATGAGATTGAA 2334  
Db 241 GACTCGGATCAATCTCAGGCTTCAATGGAATTCATCTCAAGATGAGATTGAA 300  
QY 2335 TATGAAACACGAGAGGCTGGCAGAGATTTTAAAGAAAGTAAAGA 2394  
Db 301 TATGAAACACGAGAGGCTGGCAGAGATTTTAAAGAAAGTAAAGA 354  
QY 2395 GAAGACTCTCTTCTGCTTTCAGGCTTCAAGTGGCAGAGATTTTCAAGTGGCAGAG 2454  
Db 355 GAAGATCTCTTCTGCTTTCAGGCTTCAAGTGGCAGAGATTTTCAAGTGGCAGAG 414  
QY 2455 TCCTTGATTTTGAATATTTGCTATGCTGATGCTCAACTCTCAAGTGGCAGAG 2514  
Db 415 TCCTTGATTTTGAATATTTGCTATGCTGATGCTCAACTCTCAAGTGGCAGAG 474  
QY 2515 AAGATCTGCTGCTTTCAGGCTTCAAGTGGCAGAGATTTTCAAGTGGCAGAG 2574





```
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7976
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-040-862-7976

Query Match      13.0%; Score 448.8; DB 9; Length 597;
Best Local Similarity 85.4%; Pred. No. 6.5e-113; Indels 6; Gaps 1;
Matches 514; Conservative 0; Mismatches 82;

Qy 2035 CTGGGACACCATGACATCGTGAATCTGCTGGGGCATGCACACACTGTCAGGGCCAGTG 2094
Db 1 CTGGGAAGCCACGAGAAATATTGTAACCTCTGCTGGGGCGTGCACACTGTCAGGACCAATT 60

Qy 2095 TACTTGATTTTGAATATTGTTGCTATGGTGACCTCTCAACTACCTAAGAAAGTAAAGA 2154
Db 61 TACTTGATTTTGAATATTGTTGCTATGGTGACCTCTCAACTATCTAAGAAAGTAAAGA 120

Qy 2155 GAGAAGTTTCACAGGACATGCACAGAGATTTTAAAGAACATAATTTTCAGTTCTTACCCT 2214
Db 121 GAAAAATTTTCACAGGACATGGAACCTGCTGGGGGGTGCACACTGTCAGGACCAATT 180

Qy 2215 ACTTTCAGGCACATTCAAATTCAGCATGCTGTTTCAGAGAGAGTTTCACTTACACCG 2274
Db 181 ACTTTCAGGCACATTCAAATTCAGCATGCTGTTTCAGAGAGAGTTTCACTTACACCG 240

Qy 2275 CCCTTGGATCTCAGAGACATGGAATTCCTTCAAGAGAGTTTCACTTACACCG 2334
Db 241 GACTCGGATCAAAATCTCAGGCTTCATGGGAAATTCATTTCACTTCAAGAGTAAAGA 300

Qy 2335 TATGAAACACGAGAGGCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2394
Db 301 TATGAAACACGAGAGGCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 354

Qy 2395 GAAGACCTCTTTCAGTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2454
Db 415 TCGTGTTTCACAGAGACCTGGCCCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 474

Qy 2455 AAGATCTGTGACTTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTG 2514
Db 474 AAGATCTGTGACTTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTG 474

Qy 2515 AAGATCTGTGACTTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTG 2574
Db 474 AAGATCTGTGACTTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTG 474

Qy 2575 GCAAGCAGCGCTGCGCGTGAAGTGGATGGATGGATGGATGGATGGATGGATGGATGGAT 2634
Db 535 GCAAGCAGCGCTGCGCGTGAAGTGGATGGATGGATGGATGGATGGATGGATGGATGGAT 594

Qy 2635 AC 2636
Db 595 AC 596

RESULT 14
US-09-796-692-8273
; Sequence 8273, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09796.692
; CURRENT FILING DATE: 2001-03-01
; PRIOR FILING DATE: 2001-03-01
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
```

```
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8273
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (586)
; OTHER INFORMATION: n=A,T,C or G
; US-09-796-692-8273

Query Match      13.0%; Score 448; DB 9; Length 601;
Best Local Similarity 84.8%; Pred. No. 1.1e-112;
Matches 515; Conservative 0; Mismatches 86; Indels 6; Gaps 1;

Qy 2035 CTGGGACACCATGACACATCGTGAATCTGCTGGGGCATGCACACTGTCAGGGCCAGTG 2094
Db 1 CTGGGAAGCCACGAGAAATATTGTAACCTGCTGGGGGGTGCACACTGTCAGGACCAATT 60

Qy 2095 TACTTGATTTTGAATATTGTTGCTATGGTGACCTCTCAACTACCTAAGAAAGTAAAGA 2154
Db 61 TACTTGATTTTGAATATTGTTGCTATGGTGACCTCTCAACTATCTAAGAAAGTAAAGA 120

Qy 2155 GAGAAGTTTCACAGGACATGGAACAGAGATTTTAAAGAACATAATTTTCAGTTCTTACCCT 2214
Db 121 GAAAAATTTTCACAGGACATGGAACAGAGATTTTCAAGGAACACAAATTTTCAGTTTTCACCC 180

Qy 2215 ACTTTCAGGCACATTCAAATTCAGCATGCTGTTTCAGAGAGAGTTTCACTTACACCG 2274
Db 181 ACTTTCAGGCACATTCAAATTCAGCATGCTGTTTCAGAGAGAGTTTCACTTACACCG 240

Qy 2275 CCCTTGGATCTCAGAGACATGGAATTCCTTCAAGAGAGTTTCACTTACAGATGAGTAA 2334
Db 241 GACTCGGATCAAAATCTCAGGCTTCATGGGAAATTCATTTCACTTCAAGAGTAAATGAA 300

Qy 2335 TATGAAACACGAGAGGCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2394
Db 301 TATGAAACACGAGAGGCTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 354

Qy 2395 GAAGACCTCTTTCAGTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2454
Db 355 GAAGATCTCTTTCAGTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 474

Qy 2455 AAGATCTGTGACTTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTG 2514
Db 474 AAGATCTGTGACTTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTG 474

Qy 2515 AAGATCTGTGACTTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTG 2574
Db 474 AAGATCTGTGACTTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTG 2574

Qy 2575 AAGATCTGTGACTTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTG 534
Db 475 AAGATCTGTGACTTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTGGAGTTG 534
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 03:00:43 : Search time 4534.48 seconds  
(without alignments)  
12332.831 Million cell updates/sec

Title: US-09-919-408-1  
Perfect score: 3453  
Sequence: 1 GCGGCTGGCTACCGCGC.....AAAAAAAAAAAAAAAAAAAA 3453

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estcov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_estc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_est3.\*  
12: gb\_est4.\*  
13: gb\_est5.\*  
14: gb\_estfun.\*  
15: em\_estom.\*  
16: gb\_gss.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_other.\*  
25: em\_gss\_pro.\*  
26: em\_gss\_fod.\*  
27: em\_gss\_fod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	655.8	19.0	765	9	AI323253 mp88b02.y
2	556.8	16.1	618	9	AA120050 mp88b02.r
3	531.8	15.4	604	12	BF523018 UI-R-C2p-
4	502.4	14.5	747	9	AI323643 mp88b02.x
5	486.6	14.1	522	10	BB644407 BB644407
6	452.6	13.1	621	13	BI360262 387099 MA

c	7	445	12..9	445	10	BE651447
	8	432.8	12.5	714	13	BI461248
	9	426.2	12.3	515	14	BQ556177
c	10	423	12.3	484	13	BM117950
	11	354.4	10.3	507	13	BM484050
	12	340.8	9.9	460	10	AV713950
	13	333.6	9.7	3312	11	AK004947
	14	324.8	9.4	5174	11	BC026713
	15	314	9.1	347	12	BG145808
c	16	294	8.5	457	9	AA996414
	17	277.6	8.0	357	9	AI556371
c	18	273	7.9	424	9	AI180146
	19	266	7.7	502	12	BF557871
	20	265.8	7.7	455	13	BM147250
	21	254	7.4	568	10	BE679272
c	22	248.2	7.2	375	9	AA851570
	23	244.8	7.1	660	10	BE047675
c	24	227.8	6.6	386	9	AI178876
	25	227.6	6.6	407	10	AV596131
	26	226.6	6.6	739	12	BF159205
	27	226.4	6.6	294	10	BB247268
	28	224.8	6.5	666	9	AL047744
	29	224.6	6.5	503	13	BI481351
	30	219.4	6.4	666	13	BG934591
	31	219.2	6.3	286	10	BB332140
	32	216	6.3	848	13	BI107537
	33	215	6.2	940	9	AI325368
c	34	214.4	6.2	702	12	BF163671
	35	213.4	6.2	469	12	BF189757
	36	212	6.1	696	10	BE260791
	37	212	6.1	794	13	BI111563
	38	211	6.1	832	13	BI694328
	39	209.8	6.1	531	12	BF320739
	40	207.6	6.0	575	12	BF930052
	41	205	5.9	696	14	BM951127
	42	202.8	5.9	665	13	BG922846
	43	201.4	5.8	412	12	BE853746
	44	201	5.8	290	10	BB282068
	45	199.4	5.8	669	13	BJ012204

## ALIGNMENTS

RESULT 1  
AI323253  
LOCUS  
DEFINITION  
AI323253  
mp88b02.y1 Soares.thymus.2NBMT Mus musculus cDNA clone IMAGE:576267  
5' similar to gb:U02687 FL CYTOKINE RECEPTOR PRECURSOR (HUMAN); PDGF  
gb:X59398 Mouse Flt3 mRNA for tyrosine kinase receptor of the PDGF  
(MOUSE);, mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AI323253 765 bp mRNA linear EST 23-DEC-1998  
AI323253.1 GI:4057682  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 765)  
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,  
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,  
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,  
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and  
Waterston,R.  
The WashU-HMI Mouse EST Project  
Unpublished (1996)  
Contact: Marra M/Mouse EST Project  
WashU-HMI Mouse EST Project  
Washington University School  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: mouseest@watson.wustl.edu

TITLE  
JOURNAL  
COMMENT

This clone is available royalty-free through LNLN : contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:350915  
This read is a RESEQUENCE of a previously sequenced mouse clone  
This read has been verified (found to hit its original self in the correct orientation)  
Seq primer: -40Kbp from Gibco  
High quality sequence stop: 423.

FEATURES  
SOURCE

	Location/Vxualileis	
I. .765	/organism="Mus musculus"	
	/strain="C57BL/6J"	
	/db_xref="taxon:10090"	
	/clone="IMAGE:576267"	
	/clone_lib="Soares_thymus_2NbMT"	
	/sex="male"	
	/tissue_type="Thymus"	
	/dev_stage="4 weeks"	
	/lab_host="DH10B"	
	/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTCAACTCTGAAGTGCGAGCGCGGTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."	
BASE COUNT	203 a 176 c 182 g 200 t	4 others
ORIGIN		
Query Match	19.0%; Score 655.8;	DB 9; Length 765;
Best Local Similarity	94.6%;	Pred. No. 1.1e-103;
Matches 699;	Conservative 0;	Mismatches 36; Indels 4; Gaps 2;
QY	1558 GGCCTTCGTGCTCAATGCTGTCGGTCAAAATTCATTGGCACGCTTCGCGAACCACCTCTTT	1617
Db	1 GGCGTTCTGTGCTCAATGCTGTCGGTCAAAATTCATTGGCACGCTTCGCGAACCACCTCTTT	60
QY	1618 TTAACCTCACCGAGCCGCCCTTCCCTTTTCATCCAAGAACAATCTCCTTCTATGGAGCATTT	1677
Db	61 TTAACCTCACCGAGCCGCCCTTCCCTTTTCATCCAAGAACAATCTCCTTCTATGGAGCATTT	120
QY	1678 GGGCTCTGCTCCCCCTTCATGTTGTTCTCATTTGTTGATCTGCCACAATACAAAAG	1737
Db	121 GGGCTCTGCTCCCCCTTCATTTGTTCTCATTTGTTGATCTGCCACAATACAAAAG	180
QY	1738 CAATTTAGGTACGAGATCAGTCCAGATCCAGGTGACTGGCCCTTGGATAACGAG	1797
Db	181 CAATTTAGGTACGAGATCAGTCCAGATCCAGGTGACTGGCCCTTGGATAACGAG	240
QY	1798 TACTTCTACGTTGACTTTCAGGGACTATGAATATGACCTTAAGTGGGAGTTCCCGAGAGAG	1857
Db	241 TACTTCTACGTTGACTTTCAGGGACTATGAATATGACCTTAAGTGGGAGTTCCCGAGAGAG	300
QY	1858 AACTTAGAGTTTGGGAAGGTCCTGGGCTCTGGGCTTTCGGGAGGGTGATGAACGCCACG	1917
Db	301 AACTTAGAGTTTGGGAAGGTCCTGGGCTCTGGGCTTTCGGGAGGGTGATGAACGCCACG	360
QY	1918 GCCTATGGCATTAGTAAACGGGAGTCTCAATTCAGGTGGGGTGAGATGCTTAAAAAGAG	1977
Db	361 GCCTATGGCATTAGTAAACGGGAGTCTCAATTCAGGTGGGGTGAGATGCTTAAAAAGAG	420
QY	1978 AAAGCTCACAGCTCTGAAAAGAAGCTCTCATGTGCGGAGCTCAAAATGATGACCCACCTG	2037
Db	421 AAAGCTCACAGCTCTGAAAAGAAGCTCTCATGTGCGGAGCTCAAAATGATGACCCACCTT	480
QY	2038 GGACACCATGACACATCGTGAATCTGCTGGGGCGATGCACACCTGTGAGGGCCAGTGTAC	2097
Db	481 GGACACCATGACACATCGTGGATTTGTTGGGGCTTCCACACTGCCAGGGCCAGTGCAC	540
QY	2098 TTGATTTTGAATATTGTTGCTATGGTGACCTCCTCAACTACCTTAAGAAAGTAAAAAGAGAG	2157

[illegible]

## RESULT 2

```

AA120050      618 bp      mRNA      linear      EST 17-FEB-1997
mp88502.r1 Soares.thymus_2NBMT Mus musculus cDNA clone IMAGE:576267
5' similar to gb:M64689 Mouse flk-2 mRNA, complete cds (MOUSE));
mRNA sequence.
AA120050
AA120050      1 GI:1677559
EST.
house mouse.
mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 618)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubouque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMMI Mouse EST Project
Washington University School of Medicine#
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image-llnl.gov) for further information.
MGI:350915
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop:335.

```

FEATURES  
SOURCE

```

1. .618
source
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:576267"
/clone_lib="Soares_thymus_2NbWT"
/sex="male"
/tissue_type="Thymus"
/lab_stage="4 weeks"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Site_2: Eco RI; 1st strand.cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTCACATCTCAATGGAGCGCGGTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M.Fatima Bonaloo."
158 a 142 c 158 g 160 t
BASE COUNT
ORIGIN

```

[illegible]

Query Match	16.13;	Score 556.8;	DB 9;	Length 618;
Best Local Similarity	97.7%;	Pred. No. 1.6e-86;		
Matches 586;	Conservative 0;	Mismatches 12;	Indels 2;	Gaps 2;
Qy	1558	GGGCTTCTGGTCAAATGCTGTGCGTCAAAATCTATATGGCAGCTCTTCGCGAAACCATCTTT	1617	
Db	1	GGGCTTCTGGTCAAATGCTGTGCGTCAAAATCTATATGGCAGCTCTTCGCGAAACCATCTTT	60	
Qy	1618	TTAAACTCACAGGCCCTTCCCTTTTCATCCAGACAACATCTCTTCTATGCGACCATT	1677	
Db	61	TTAAACTCACAGGCCCTTCCCTTTTCATCCAGACAACATCTCTTCTATGCGACCATT	120	
Qy	1678	GGGCTCTGCTCCCTTTCATGTGTGTTCTCATTCTGTGTGATCTGCCACAATACAAAAG	1737	
Db	121	GGGCTCTGCTCCCTTTCATGTGTGTTCTCATTCTGTGTGATCTGCCACAATACAAAAG	180	
Qy	1738	CAATTTAGGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCTCGGATAACGAG	1797	
Db	181	CAATTTAGGTACGAGAGTCAGCTGCAGATGATCCAGGTGACTGGCCCTCGGATAACGAG	240	
Qy	1798	TACTTCTACGTTGACNTCAGGGACTATGAATATGACCTTAAGTGGGAGTTCCCGAGAGAG	1857	
Db	241	TACTTCTACGTTGACNTCAGGGACTATGAATATGACCTTAAGTGGGAGTTCCCGAGAGAG	300	
Qy	1858	AACTTAGAGTTTGGGAAGGTCCTGGGCTCTGGGCTTTCGSGAGGGTGATGAAGCCACG	1917	
Db	301	AACTTAGAGTTTGGGAAGGTCCTGGGCTCTGGGCTTTCGSGAGGGTGATGAAGCCACG	360	
Qy	1918	GCTATGGCATTAGTAAACGGGAGTCTCAATTGAGTGGCGGTGAAGATGCTAAAGAG	1977	
Db	361	GCTATGGCATTAGTAAACGGGAGTCTCAATTGAGTGGCGGTGAAGATGCTAAAGAG	420	
Qy	1978	APAGCTGCAGAGCTGTGAARAAGACTCTCATGTCGGAGCTCAAAATGATGACCCACCTG	2037	
Db	421	APAGCTGCAGAGCTGTGAARAAGACTCTCATGTCGGAGCTCAAAATGATGACCCACCTG	480	
Qy	2038	GGACACCATGACAACATCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCCAGTGTAC	2097	
Db	481	GGACACCATGACAACATCGTGAATCTGCTGGGGCATGCACACTGTGAGGGCCAGTGTAC	540	
Qy	2098	TTGATTTTTGAATATTTGTTGCTATGTTGACCTCCTCACTACCTAGAAAGTAAAGAGAG	2157	
Db	541	CTGATTTTTGAATATTTGTTGCTATGTTGACCTCCTCACTACCTAGAAAGTAAAGAGAG	598	

RESULT 3	BF523018	604 bp	linear	EST 11-DEC-2000
LOCUS	UI-R-C2p-rg-f-10-0-UI.r1	UI-R-C2p	Rattus norvegicus	CDNA clone
DEFINITION	UI-R-C2p-rg-f-10-0-UI 5',	RNA sequence.		
ACCESSION	BF523018			
VERSION	BF523018.1	GI:11631033		
KEYWORDS	EST.			
SOURCE	Norway rat.			
ORGANISM	Rattus norvegicus			
REFERENCE	1 (bases 1 to 604)			
AUTHORS	Ronald M.F., Lennon G. and Soares M.B.			
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery			
JOURNAL	Genome Res. 6 (9), 791-806			
MEDLINE	9704477			
COMMENT	Contact: Soares, MB			
	Program for Rat Gene Discovery and Mapping			
	University of Iowa			
	451 Eckstein Medical			
	Research Building Iowa City, IA 52242, USA			
	Tel: 319 335 8250			
	Fax: 319 335 9565			
	Email: msoares@blue.weeg.uiowa.edu			
	CDNA Library Preparation: M.B. Soares Lab			
	Clones will be available through Research Genetics (www.resgen.com)			

```

This clone is also available through the I.M.A.G.E. Consortium at
LNLN (info@image.llnl.gov). IMAGE ID= 1793119
Seq primer: M13 Forward.
Location/Qualifiers
1. .604
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C2p-rg-f-10-0-UI"
/clone_lib="UI-R-C2p"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-C2p
library is a subtracted library derived from the UI-R-C1
library, which is a subtracted library derived from the
UI-R-C0 library. The UI-R-C0 library consisted of a
mixture of individually tagged normalized libraries
constructed from rat placenta, adult lung, brain, liver,
kidney, heart, spleen, ovary, muscle, 8', 12 and 18-day
embryo. The tag is a string of 3-5 nucleotides present
between the Not I site and the oligo-dT track which allows
identification of the library of origin of a clone within
the mixture. The subtracted library (UI-R-C2p) was
constructed as follows: PCR amplified cDNA inserts from
UI-R-C1 clones from which 3' ESTs had been derived was
used as a driver in a hybridization with the UI-R-C1
library in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library) was
purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
UI-R-C2p library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research 6:
791-806, 1995)."

```

[illegible]



51 CGGCTGCTGCTGCTTGTGTTTGTGTCAGTAATGATCTTGAGACCGTTACAAACCAAGAC 120

FEATURES	Location/Qualifiers
source	1..621

```

FEATURES
source
Location/Qualifiers
1. .621
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DHI08"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

```



```
BASE COUNT      190 a      132 c      141 g      158 t
ORIGIN

Query Match      13.1%; Score 452.6; DB 13; Length 621;
Best Local Similarity 83.2%; Pred. No. 1.6e-68;
Matches 515; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 1702 GTTCTCATTTGCTGATCGGCACAAATACAAAAGCAATTTAGTACGAGAGTGCAGTGT 1761
DB 3 GTTTTAAACATGCTTAATTTGTCACAAAGTACAAAAGCGATTCCGTCAGAAAGCCAGCTG 62

QY 1762 CAGATGATCCAGGTGCTGGCCCTGGATACGAGTACTTCTAGCTTGCATTCAGGAC 1821
DB 63 CAGATGGTCCAGGTGACGGGTCCTGGTACGACTACTTCTATCAGCTTCAGAGAA 122

QY 1822 TATGAATATGACCTTAAGTGGGAGTTCGCCGAGAGAACTTATAGATTTGGGAAGTCTCTG 1881
DB 123 TATGAATATGACCTCAATTTGGAGTTTCCAGGAAAAATTTAGAAATTTGGGAAGTCTCTG 182

QY 1882 GGGTCTGGGCTTTCCGAGGGTGTATGACGCCAGCGCTATGGCATTTAGTAAACGGGA 1941
DB 183 GGATCCGGCGCTTTTGGAAAAAGTGTAAACGCACTGCTATGGAAATCAGTAAACTGGA 242

QY 1942 GTCTCAATTCAGTGGCGTGAAGATGCTAAAAGAGAAAGCTGCAGCTGTGAAAAAGAA 2001
DB 243 GTGTCAATCCAGTGGCAGTCAAAATGCTGAAGAAGAAAGCGACAGCTCGGAACGAGAG 302

QY 2002 GCTCTCATGTCGGAGCTCAAAATGATGACCCACCTGGGACACCATGACAACTCGTGAAT 2061
DB 303 GCTCTCATGTCGTAATCAAAATGATGACCCACCTGGGACGAGCAAGAAATAGTGAAC 362

QY 2062 CTGCTGGGGCATGACACTGTCAGGGCCAGTGTACTTCAATTTTCAATTTTGTGCTAT 2121
DB 363 CTGCTGGGGCATGACACTGTCAGGGCCAGTGTACTTCAATTTTGAATTTTGTGCTAC 422

QY 2122 GTGTGACCTCCTCACTTACCTAAGAAAGTAAAGAGAAAGTTCACAGGACATGGACAGAG 2181
DB 423 GTGTGATCTCTCACTATCTAGAGTAAAGAGAAATTCATAGGACATGGACCGAG 482

QY 2182 ATTTTAAAGAACATAATTTCACTTCTTACCTTCTTCCAGGACATTCATTAATCCAGC 2241
DB 483 ATTTTCAAGGAACATAATTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 542

QY 2242 ATGCTCTGTTTCAGGAAGTTCAGTACACCCGCTTGGATCAGCTCTCAGGGTTCAT 2301
DB 543 ATGCTCTGTTTCAGGAAGTTCATTAATACACCCAGACTCGGATCTTCTTCTTCTTCTTCT 602

QY 2302 GGAATTTCAATTCATCTG 2320
DB 603 GGAATTCATTCATCTG 621

RESULT 7
BE651447/c      445 bp      mRNA      linear      EST 06-SEP-2000
LOCUS      UI-M-BH3-atk-d-06-0-UI.r1 NIH_BMAP_M_S4 Mus musculus cDNA clone
DEFINITION      UI-M-BH3-atk-d-06-0-UI 5', mRNA sequence.
ACCESSION      BE651447
VERSION        BE651447
KEYWORDS       EST.
SOURCE         house mouse.
ORGANISM       Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 445)
AUTHORS        Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE          Normalization and subtraction: two approaches to facilitate gene
               discovery
JOURNAL        Genome Res. 6 (9), 791-806 (1996)
MEDLINE        97044477
COMMENT        Contact: Chin, H
               National Institute of Mental Health
               6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
```

```
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: meste@mail.nih.gov
CDNA Library Preparation: M.B. Soares Lab Clone distribution:
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
should be noted that Bento Soares is generating a small number of
additional specialized non-redundant arrays of BMAP cDNAs whose
availability will be considered under appropriate and limited
collaborative arrangements
Seq primer: M13 Reverse.
Location/Qualifiers
1. .445
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH3-atk-d-06-0-UI"
/clone_lib="NIH_BMAP_M_S4"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not 1; Site_2: Eco RI; The
NIH_BMAP_M_S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH_BMAP_M_S4,
NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
(NIH_BMAP_M_S4) was constructed as follows: PCR amplified
cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)
was purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
NIH_BMAP_M_S4 library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)."
6:791-806, 111 g 115 t
BASE COUNT      96 a      123 c      111 g      115 t
ORIGIN

Query Match      12.9%; Score 445; DB 10; Length 445;
Best Local Similarity 100.0%; Pred. No. 3.7e-67;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 612 TGTGGAGTGGGTCTCTGCAGCTCCACAGGGAAGCTGTAAAGAAAGAGCCCTGTGTGT 671
DB 445 TGTGGAGTGGGTCTCTGCAGCTCCACAGGGAAGCTGTAAAGAAAGAGCCCTGTGTGT 386

QY 672 TGTCAAGAGGAGGAAAAGTACTTATGATGTTTTCGGACACAGACATCATCTGTC 731
DB 385 TGTCAAGAGGAGGAAAAGTACTTATGATGTTTTCGGACACAGACATCATCTGTC 326

QY 732 TAGAAATGCACTGGCGCGCAATGCACCAAGCTGTTCCACCATGATCTAAACACCGCTCC 791
DB 325 TAGAAATGCACTGGCGCGCAATGCACCAAGCTGTTCCACCATGATCTAAACACCGCTCC 266

QY 792 TCAGACACACTGCCCCAGTATTCTCTGAAAGTGGGGAACCCCTTGTGGATCAGGTGTA 851
DB 265 TCAGACACACTGCCCCAGTATTCTCTGAAAGTGGGGAACCCCTTGTGGATCAGGTGTA 206

QY 852 GGCCATTCATGTGAACCATGATTCGGGCTCACCCTGGGAGCTGGAACAGAACCCCTGGA 911
DB 205 GGCCATTCATGTGAACCATGATTCGGGCTCACCCTGGGAGCTGGAACAGAACCCCTGGA 146
```

```

QY 912 GGAGGGCAGCTACTTTGAGATGAGTACCTACTCCACAACAGGACCATGATTCGGATTCT 971
Db 145 GGAGGGCAGCTACTTTGAGATGAGTACCTACTCCACAACAGGACCATGATTCGGATTCT 86
QY 972 CTTGGCCTTTGTGCTTCCTGGGGAAGAACAGACACCGGATATTACACTGCTCTTCCTC 1031
Db 85 CTTGGCCTTTGTGCTTCCTGGGGAAGAACAGACACCGGATATTACACTGCTCTTCCTC 26
QY 1032 AAAGCACCCAGCCAGTCAGCGTTG 1056
Db 25 AAAGCACCCAGCCAGTCAGCGTTG 1

RESULT 8
LOCUS BI461248 714 bp mRNA linear EST 21-AUG-2001
DEFINITION 603206574F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:5272266 5',
mRNA sequence.
ACCESSION BI461248
VERSION BI461248.1 GI:15251904
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 714)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L14M11687 row: c column: 19
High quality sequence stop: 709.
Location/Qualifiers
1. 714
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5272266"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTTNN-3',
size-selected for average insert size 2.2 kb and
normalized to ROT 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 182 a 171 c 191 g 170 t
ORIGIN
Query Match 12.5%; Score 432.8; DB 13; Length 714;
Best Local Similarity 79.5%; Pred No. 4e-65;
Matches 525; Conservative 0; Mismatches 132; Indels 3; Gaps 1;
QY 1 CGCGCCTTGGCTACCGCGCTCCGAGGCCATCGGGGTGGGCGCAGCGAGCCGCG 60
Db 58 CGCGCCTGGGACCCCGGCTCCGAGGCCATCGGGGTGGC---CGCGCAGCGCGC 114
QY 61 CGGCTGCTGCTGCTGTTGTTTGTCAATGATCTTGAGACCGTTTACAAACAGAC 120
Db 115 CAGCTGCGGCTGCTGCTGTTTCTGCAATGATATTGGGACTATTACAAATCAAGAT 174
QY 121 CTGCTGTGATCAAGTGTGTTTAAATCAGTCATCAGAACAAATGGCTCATCAGCGGGAAG 180

```

```

Db 175 CTGCGCTGTGATCAAGTGTGTTTAAATCAATCATAGAACAATGATTCATCAGTGGGAAG 234
QY 181 CCATCATGCTACCGAATGTCGAGGATCCCGAGAAGACCTCCAGTGTACCCCGAGCGGC 240
Db 235 TCATCATATATATCCCATGATATCAGAAATCCCGGAAGACCTCGGGTGTGCTGTGAGACCC 294
QY 241 CAGAGTGAAGGACGATATATGAAGCGCCACCGTGGAGTGGCCGAGTCTGGTCCATC 300
Db 295 CAGAGCTCAGGACAGTGTACGAAGCTGCCGCTGTGGAAGTGGATGTATCTGCTTCCATC 354
QY 301 ACCCTGCAAGTGCAGCTGCCACCCAGGGGACCTTTCTCTGCTCTGGGTCTTTAAGCAC 360
Db 355 ACACCTGCAAGTGCAGTGCAGTGCAGGAGCAATTTCTCTGCTCTGGGTCTTTAAGCAC 414
QY 361 AGCTCCCTGGGTCGTCAGCGGCACCTTTGATTTACAAAACAGAGGAATCGTTTCCATGGCC 420
Db 415 AGCTCCCTGAAATGTCAGCGGCACATTTGATTTACAAAACAGAGGAGTAGTTTCCATGGTC 474
QY 421 ATCTTTGAAGCTGCAGAGACCCAGGAGGAGTACCTACTCCATATTCAGAGCGAACGC 480
Db 475 ATTTTGAAGTGCAGAAACCCAGCTGGAGATACCTACTTTTATTCAGAGTGAAGCT 534
QY 481 GCCAACTACACAGTACTGTTTACAGTGAATGTAAGAGATACACAGCTGTATGCTAAGG 540
Db 535 ACCAATTACACAATATGTTTACAGTGAGTATAAGAAATACCTGCTTTACACATTAAGA 594
QY 541 AGACCTTACTTTAGGAAGATGAAACACAGGATGCACTGCTGTCGATCTCCGAGGGTGT 600
Db 595 AGACCTTACTTTAGAAAATGAAACACAGGAGCGCTGCTGTCATATCTGAGAGCGTT 654
QY 601 CGGAGGCCACCTGTGGAGTGGCTGCTGTCAGCTCCACAGGAGAAAGCTGTAAGAGAA 660
Db 655 CCAGAGCCGCTGCTGGAATGGGTGCTTTCGATTCACAGGGGGAAGCTGTAAGAGAA 714

RESULT 9
LOCUS BQ556177 515 bp mRNA linear EST 20-JUN-2002
DEFINITION H4038E08-5 NIA Mouse 7.4K cDNA clone Set Mus musculus cDNA clone
ACCESSION BQ556177
VERSION BQ556177.1 GI:21457065
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 515)
VanBuren, V., Piao, Y., Dudekula, D.B., Qian, Y., Carter, M.G., Martin
P.K., Stagg, C.A., Bassey, O., Alpa, K., Hamatani, T., Kargul, G.J.,
Luo, A.G. and Ko, M.S.H.
Assembly, verification, and initial annotation of NIA 7.4K mouse
cDNA clone set
Unpublished (2002)
Other_ESTS: H4038E08-3
Contact: Rong Qian
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit http://lgsun.grc.nia.nih.gov/cDNA/NIA_7.4k.html for details.
Plate: H4038 row: E column: 08
Seq primer: -21M13 Reverse
High quality sequence stop: 515
POLYA-No.
Location/Qualifiers
1. 515
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="niaEST:H4038E08-5"
/db_xref="taxon:10090"

```

FEATURES  
source

```

/clones="H4038E08"
/clone_lib="NIA Mouse 7.4K cDNA Clone Set"
/sex="mixed"
/dev_stage="mixed"
/lab_host="DH10B"
/notes="vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
clone is among a rearranged set of 7,407 clones from more
than 20 cDNA libraries."
BASE COUNT      139 a 138 c 128 g 110 t
ORIGIN
Query Match      12.3%; Score 426.2; DB 14; Length 515;
Best Local Similarity 92.8%; Pred. NO. 6.2e-64;
Matches 478; Conservative 0; Mismatches 13; Indels 24; Gaps 2;

Qy 2725 GCTAACTCTATAAAGTATGATTCAGAGTGGATTTAAATGGAGCGCATTTCTATGCCACA 2784
Db 1 GCTAACTCTATAAAGTATGATTCAGAGTGGATTTAAATGGAGCGCATTTCTATGCCACA 60

Qy 2785 GAAGGATATATCTTTGTAATCAATCCTGCTGGGCTTTTGACTCAAGGAAGCGGCCATCC 2844
Db 61 GAAGGATATATCTTTGTAATCAATCCTGCTGGGCTTTTGACTCAAGGAAGCGGCCATCC 120

Qy 2845 TTCCCAACCTGACTTCATTTTATAGATGTCAGCTGGCAGGCGGAGAAGC----- 2897
Db 121 TTCCCAACCTGACTTCATTTTATAGATGTCAGCTGGCAGGCGGAGAAGCGCATGTAT 180

Qy 2898 -----ATGATCAGACATCCATCCATCTACCAAAACAGCGGCCCTC 2941
Db 181 CAGAACATGGGTGCAACGTCCTCCAGACATCCATCCATCTACCAAAACAGCGGCCCTC 240

Qy 2942 AGCAGAGGCGGCTCAGAGCC-CAGTCGCCACAGCGCCAGGTGAAGATTCAAGAGAA 3000
Db 241 AGCAGAGGCGGCTCAGAGCCGCGCATCGCCACAGGCGCCAGGTGAAGATTCAAGAGAA 300

Qy 3001 AGAAGTTAGCGAGGAGGCTTTGGACCCCGCCACCCCTAGCAGGCTGTAGACCGCAGAGCA 3060
Db 301 AGAAGTTAGCGAGGAGGCTTTGGACCCCGCCACCCCTAGCAGGCTGTAGACCGCAGAGCA 360

Qy 3061 AGATTAGCTCGCTCTCAGGAGCGCCCTACAGCGCGTGTCTTCGCTGGACTTTTCTCT 3120
Db 361 AGATTAGCTCGCTCTCAGGAGCGCCCTACAGCGCGTGTCTTCGCTGGACTTTTCTCT 420

Qy 3121 AGATCTCTCTGCCATTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCTCGCACAG 3180
Db 421 AGATCTCTCTGCCATTACTCCAAAGTGACTTCTATAAATCAAACTCTCTCTCGCACAG 480

Qy 3181 GCGGAGAGGCCAATAATGAGACTTGTGGTGAGCC 3215
Db 481 GTGGGAGAGCCAAATAATGAGACTTGTGGTGAGCC 515

RESULT 10
BM117950/c
LOCUS      484 bp mRNA linear EST 30-JAN-2002
DEFINITION L0858E03-3 NIA Mouse Newborn Brain cDNA Library Mus musculus cDNA
clone L0858E03 3', mRNA sequence.
ACCESSION BM117950
VERSION    1
KEYWORDS  EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerogathia; Muridae; Murinae; Mus.
1 (bases 1 to 484)
Pao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Pantano, S., Lim, M.K.
and K.O.M.S.H.
Systematic Analyses of NIA Mouse Newborn Brain cDNA Library
Unpublished (2001)
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

```

Email: cdna@lgsun.grc.nia.nih.gov  
 Plate: L0858 row: E column: 03  
 Seq primer: -21M13 Forward  
 High quality sequence stop: 484  
 POLYA=Yes.

#### FEATURES

Location/Qualifiers  
 1..484  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="niaEST:L0858E03-3"  
 /db\_xref="taxon:10090"  
 /clone\_lib="NIA Mouse Newborn Brain cDNA Library"  
 /tissue\_type="Newborn Brain"  
 /dev\_stage="Newborn"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pSPORT1 (Invitrogen); Site\_1:  
 SalI; Site\_2: NotI; Mouse cDNA project by the Laboratory  
 of Genetics, National Institute on Aging (NIA), Intramural  
 Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).  
 Double-stranded cDNAs were synthesized with an Oligo(dT)  
 primer (Invitrogen: 5'-  
 PGACTAGTCTAGATCGGAGCGCGCCCTTTT-3') from 48  
 microgram of total RNA, treated with T4 DNA polymerase,  
 and purified by ethanol-precipitation. The cDNAs were  
 ligated to Loner-linker LL-Sal3 (Ref. Development  
 127:1737-1749 (2000) [PMID:10725249]), purified by  
 phenol/chloroform, and separated from free linkers by  
 Centricon 100. Then, the cDNAs were digested with SalI and  
 NotI enzymes, and cloned into SalI/NotI site of pSPORT1  
 plasmid vector. The DH10B E. Coli host was transformed  
 with ligation mixture by the standard chemical method. The  
 average insert size is about 1.9 kb. The library was  
 constructed by Yulan Piao (NIA)."

BASE COUNT 127 a 109 c 114 g 134 t  
 ORIGIN  
 Query Match 12.3%; Score 423; DB 13; Length 484;  
 Best Local Similarity 95.1%; Pred. NO. 2.3e-63;  
 Matches 450; Conservative 0; Mismatches 15; Indels 8; Gaps 1;  
 Qy 2963 CCCAGTCCACAGCCCGCAGGTTCAGATTTCACAGAGAAAGAGTTAGCGAGGAGGCTTG 3022  
 Db 474 CCCATCGCCACAGGCCCGCAGGTTCAGATTTCACAGAGAAAGAGTTAGCGAGGAGGCTTG 415  
 Qy 3023 GACCCGCGCACCCCTAGCAGGCTGTAGACGCGCAGAGCAAGATTAGCCTCGCCTCTCAGGA 3082  
 Db 414 GACCCGCGCACCCCTAGCAGGCTGTAGACCCACAGAGATTAGCCTCGCCTCTCAGGA 355  
 Qy 3083 AGCGCCCTACAGCGGCTTGCTTCGCTGGACTTTTCTCTAGATGCTGTCTGCCATTACTCC 3142  
 Db 354 AGCGCCCTACAGCGGCTTGCTTCGCTGGACTTTTCTCTAGATGCTGTCTGCCATTACTCC 295  
 Qy 3143 AAGTGAATCTTATAAATCAACACCTCTCTCCACAGCGGCGGAGAGCCCAATATGAGAC 3202  
 Db 294 AAGTGAATCTTATAAATCAACACCTCTCTCCACAGCGGCGGAGAGCCCAATATGAGAC 235  
 Qy 3203 TTGTTGCTGAGCGCGCTACCTCGCTGGGCGCTTT-----CCACGAGCTTGAGGGGAAA 3254  
 Db 234 TTGTTGCTGAGCGCGCTACCTCGCTGGGCGCTTTCCAGGCGCCCGCAGGCTTGAGGGGAAA 175  
 Qy 3255 GCCATCTATCTGAATATAGTATATTTCTGTAATAGCTGTAACCAACCAACCCGTTTT 3314  
 Db 174 GCCATCTATCTGAATATAGTATATTTCTGTAATAGCTGTAACCAACCAACCCGTTTT 115  
 Qy 3315 TTGCTAAGGAAAGCTAAATATGATTTTAAATCTATCTTTTAAATCTATGTAACT 3374  
 Db 114 TTGCTAAGGAAAGCTAAATATGATTTTAAATCTATCTTTTAAATCTATGTAACT 55  
 Qy 3375 TTTTCATCTATTAGTGATATATTTTATGATGGAATAAACTTTCTACTGTA 3427  
 Db 54 TTTTCATCTATTAGTGATATATTTTATGATGGAATAAACTTTCTACTGTA 2

```

RESULT 11
BM484050
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
AUTHORS
Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
TITLE
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL
COMMENT
Contact: Smith TPL
Unpublished (2000)
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCCAGTCACGACG
Plate: 2 row: D column: 23
Seq primer: ATTAGGTGACACTATAG.
FEATURES
source
1..507
Location/Qualifiers
1..507
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2Pig"
/tissue_type="pooled"
/lab_host="DRI108"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT
144 a 132 c 118 g 112 t 1 others
Query Match
Best Local Similarity 10.3%; Score 354.4; DB 13; Length 507;
Matches 425; Conservative 0; Mismatches 81; Indels 2; Gaps 2;
QY 1170 CCGATGCAGTGGATCTTCTCTCAAGCCTCATTTCTTGATGACAGAGCGCTGGAGGA 1229
D 1 CAGGTGCAGTGGACCTTCTCTC-GCAACATTTCTTGATGACAGAGCGCTGGAGGA 59
QY 1230 TGGGTACAGCATATCTAAATTTTGGCATCATAGAACAGCCAGGAGTACATATTCTA 1289
D 60 CGGTACAGCATATCTAAGTTTGGACCATTAACACCCAGGAGCATACATCTTCCA 119
QY 1290 TGCA-GAAATGATGACGCCAGCTTCCACCAAAATGTTTCAGCTGATATAGAAAGAAC 1348
D 120 TGCAGAAAATGACGAGCCAGCTTCCACAAAGATGTTTCAGCTGAATATAAGAGAAC 179
QY 1349 CTCAGTGTAGCAATGCTCAGCAGCCAGCGCTCTTCTCTGATGGGTACCCGC 1408
D 180 CGCAAGTGTGGCGAGGATCGCAAGTCAGGCTTCTCTCTCTCTGATGGGTACCCGT 239
QY 1409 TACCTCTTGGACCTGGAAGAGTGTTCGACAAATCTCCCAATTCACCGAGGAATCC 1468
D 240 TACCGTCTTGGACCTGGAAGAGTGTTCAGCAAGTCTCCCACTGCACGGAAGATCA 299
QY 1469 CACAAGGAGTTTGAATATAAAGGCTTAACAGAAAGTGTTCGCGAGTGGGTGTCGAGCA 1528
D 300 CGGAAGGCATCTGGAATATAAAGGCCAACAGGAAGATTTGGACAGTGGATTTCCAGCA 359

```

```

QY 1529 GTACTCTAAATATAGTGGAGCGGGAAGGGCTTCTGGTCAAAATGCTGTGTGTACAATT 1588
D 360 GCATCTGACATGAGGAGGCTGTCAAGGGTTTCTGGTGAAGTGTGTGCCTACAATT 419
QY 1589 CTATGGCAGGTCTTGGGAACACCATCTTTTAAACTCACCAAGGCCCCCTTCCTTCATCC 1648
D 420 CCCTGGGCAGTCTTGTGAACGATCTCTTAACTCACCAGGCCCTTCCTTCATCC 479
QY 1649 AAGACACATCTCTCTTCTATGCGACCAT 1676
D 480 AGGACACATCTCTTCTATGCAACAAT 507
RESULT 12
AV713950
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
Cheng,Z., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu
,G., Cheng,Z. and Han,Z.
TITLE
Homo sapiens cDNA DCB clones
JOURNAL
COMMENT
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzeg@sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1..460
Location/Qualifiers
1..460
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="DCB"
/cell_type="dendritic cells"
/dev_stage="mature"
/lab_host="BM25.8"
/note="vector: pRiplex2; Site_1: SfiI; Site_2: SfiI"
BASE COUNT
152 a 95 c 96 g 117 t
Query Match
Best Local Similarity 9.9%; Score 340.8; DB 10; Length 450;
Matches 378; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 1906 ATGAACCCCGCGCTATGGCATTAGTAAACAGGAGTCTCAATTCAGGTGCGGTGAAG 1965
D 1 ATGAACCCCAACAGCTTATGGAATTAGCATAACAGGTGTCTCAATCCAGTGTCCGCTCAA 60
QY 1966 ATGCTAAAGAGAAGCTGACAGCTGTGAAAGAGAGCTCTCATGTGGAGCTCAAAATG 2025
D 61 ATGCTGAAGAAAAGAGAGAGAGCTCTCAAGAGAGGACATCATGTCAAGACTCAAGATG 120
QY 2026 ATGACCCACCTGGGACACCATGACAACATCTGATCTGTGGGGGATGACACTGTCA 2085
D 121 ATGACCCAGCTGGGAGCCAGGAGATATTGTGAACCTGTGGGGGCGTGCACACTGTCA 180
QY 2086 GGGCGACTGTACTTGATTTTGAATATTGTGCTATGCTGACCTCTCAACTACCTAAGA 2145
D 181 GGACCAATTTACTTGATTTTGAATATTGTGCTATGCTGATCTCTCAACTATCTAAGA 240

```



GFIRKAVLDSNTYVCKTMVNGRESTSTGIWLKVRNVHPPEPQIKLEPSKLVIRGE  
AAQIVCSATNAEVEFNILKRGDTKLEIPLNSDFDQDNYYKKVRALSILNADVFODAGIV  
SCVASNDVGRTPATMNFQVVESEYALNTSEOSLLOEYVSGDSLILTVHADAYPSIOHY  
NWYLGPFEDORKLEITQRAIYRYTEKFLFLNRVKASEAGQYFLMAQNKAGWNLLTF  
ELTLRYPEVSVMWPNVSGDVLFCDDVSGVQPPSVTWMECHRGHTRDDEAOLQVND  
THEPLVSKRPFVKVLIQSOPLPIGLTKHNTYFCFTHNSVGNSSQYFRAVLSGQSKOLP  
DESFTFVVVACHMSVNSLLVLLLLLYKYKQPKYQVWRKIIERYEGNSYTFIDPQ  
LPYNKEFFPRNNLOFKTLGAGAFKVEATEFLGKEDAVLKVAVKMLKSTAHAD  
KEALMSLKTMISHQHENIVNLGACTHGGPVLVITEYCCYGLLFLREKAEMLG  
PSLSPGDSGDSYKHLEKKYVRDRDSGFSSQVDTYVEMRPVSTSSSFPKQDL  
DKESRPLEWDLHFFSSQAQMAFLASKNCIHRDVAARNVLTSSGVAKIGDFGLA  
RDIMNDSYVVKGNRYPVKNWAPESIFDCVYTVOSDWSYGLLWEIFSLGLNYPG  
ILVNNKFYKIDGYQMAQPVFAPKNIIYSIMQSCWDLPTRPFTFOQICFLQEQARL  
ERRDODYANLPSGSGSDSGSGSGSSSEPEESSEHLACCEPDIAOPLQPN  
NYQFC"

BASE COUNT 779 a 915 c 897 g 721 t  
ORIGIN

Query Match 9.7%; Score 333.6; DB 11; Length 3312;  
Best Local Similarity 57.8%; Pred. No. 3e-48;  
Matches 666; Conservative 0; Mismatches 459; Indels 27; Gaps 3;

QY 1697 TTGTGTTCTCATTTGTGTGCTGCCACAAATACAAAAAGCAATTTAGGTACAGAGTC 1756  
DB 1672 TGGTGCTACTGCTGCTGCTGTGTACAAAGTACAGCAAGCCGGAAGTACCAGTGC 1731  
QY 1757 AGCTGCAGATCATCCAGTGTAGTCCGCCCTGGATAACGAGTACTTCTACCTTGACTTCA 1816  
DB 1732 GCTGGAAGATCATCGAG-----AGATGGAAGCAATAGTCTACACCTTCATTGACCCTA 1785  
QY 1817 GGGACTATGATATGACCTTAAAGTGGGAGTTCCCGGAGAGAGAAGCTTAGAGTTGGGAAG 1876  
DB 1786 CTCAGTTGCCCTACAAATGAGAAGTGGGAGTTCCCTCGGAACAACCTCAGTTTGGTARA 1845  
QY 1877 TCCGTGGGCTGCGGCTTTTCGGGAGGTTGATGAACGCCACCGCTATGGCAATAGTAAAA 1936  
DB 1846 CTCTAGGAGCGGTGCTCTTTGGGAGGTTGGTGGAGGCTTACAGCCTTTGGTCTGGGCAAG 1905  
QY 1937 CGGAGTCTCAATTCAGTGGCGTGAAGTGTCTAAAGAGAAAGCTGACAGCTGTGAA 1996  
DB 1906 AAGATGAGTGTGCTGAAGGTGCTGTGAAGTGTCTAAAGTCTCCAGGCTCATCTGATGAGA 1965  
QY 1997 AAGAGTCTCATGTCGGAGCTCAAAATGATGACCCACCTCGGACACCATGACAAACATCG 2056  
DB 1966 AGGAGGCCCTCATCTCAGAGCTGAGATCATGATGATCCTCGGACACGACAGATATAG 2025  
QY 2057 TGAATCTGCTGGGGCATGACATGTCAGGGCCAGTGTACTTGATTTTGAATATTGTT 2116  
DB 2026 TCAACCTCTTTGGAGCGCTGTAATCAGCGAGGACCTGCTGCTGTCATCACTCAATCTGCT 2085  
QY 2117 GCTATGTTGACCTCTCACTCACTAAGAAAGTAAAGAGAGAAGTTTCACAGGACATGGA 2176  
DB 2086 GCTATGAGACCTTACTCAACTTTCTCGAAGAAAGCGGAGGCTATGCTAGGACCCAGCC 2145  
QY 2177 CAGAGATTTTAAAGAAATATATTTTCACTTCTTACCCTTACTTCCAGGACCATCAAAAT 2236  
DB 2146 TGAGTCTGTCAGA-----CTCCGAGGAGACTCCACCTACAGAAATCAACCTG 2198  
QY 2237 CCAGCATGCTCTGTTTCAGAGAAGTTTCACTTACACCCGCCCTTGGATCAGCTCTCAGGCT 2296  
DB 2199 GAGAAGAAATATGTGCGCAGGACAGTGGCTTCTCCAGTCAAGGTTGTAGACACCTACGTG 2258  
QY 2297 TCAATGGGAATTCATTTCTGAAGATGAGATTGAATATGAAACACCAAGAGGCTGG 2356  
DB 2259 GAGATGAGGCTGTCTCGACTTCTTCAAGTGAAGTCTCTTTTAA-----GC 2304  
QY 2357 CAGAGAAGAGAGGAGATTGTAACGCTGCTGAGCTTTGAGAGCTCTCTTTGCTTTCGCT 2416  
DB 2305 AAGATCTGGCAAGAGAGCCAGCGGCCCTTGGAGCTCTGAGGACCTGCTCCACCTTCTCCA 2364  
QY 2417 ACCAAGTGGCCAAAGGATGAAATTCCTGGAGTTCAAGTCTGCTGTCACAGAGACCTGG 2476  
DB 2365 GCCAAGTGGCTCAGGGCATGCGCTTCTGCTTCTTAAACATGCTATCCACCGGACGTAG 2424

QY 2477 CAGCCAGCAATGCTGTTGTCACCCACGGGAAGGTGTGAAGATCTGTGACTTTGGAGTCG 2536  
DB 2425 CAGCTCGAAACGTGCTGTTGACCGAGGCACATGTGCCAAGATTGGGGACTTTGGACTGG 2484  
QY 2537 CCGGAGACATCTCTGAGCGACTCCAGCTAGCTCGTCAGGGGCAACGCACGCGCTGCCGTGA 2596  
DB 2485 CTAGGACATCATGAATGACTCCCACTATGTTGTCAAGGCAATGCCCCGCTGCCTGTAA 2544  
QY 2597 AGTGGATGGCACCCGAGAGAGCTTATTTGAAGGATCTACACAATCAAGAGTGAGCTGCT 2656  
DB 2545 AGTGGATGGCCCCAGAGAGACATCTTTGACTGCGCTACACAGATTGAGAGTGATGTGCT 2604  
QY 2657 CCTAGCGCATCTCTCTCTGGGAGATATTTTCACTGGGTGAACCTTACCTGGGCATTTC 2716  
DB 2605 CCTAGCGCATCTCTCTCTGGGAGATCTTCTCGCTGGTCTGAACCTTACCCCGGATCC 2664  
QY 2717 CTGTGACGCTAACTTCTATAACTGATTCAGAGTGGATTTAAAATGGAGAGCCATTCT 2776  
DB 2665 TAGTGAACAACAAGTTCTACAACTGGTGAAGGATGGATACCAAAATGGCCAGCTGTAT 2724  
QY 2777 ATGCCACAGAGGGATATACCTTTGTAAGCAATCTCTGCTGGGCTTTTGACTCAAGAGC 2836  
DB 2725 TTGCACCGAAGAACAATATACAGCATCATGCTGCTGGGACCTGGAGCCCTACCGAA 2784  
QY 2837 GGCCATCTCTCC 2848  
DB 2785 GACCACCTTCC 2796

## RESULT 14

BC026713  
LOCUS BC026713 5174 bp mRNA linear HTC 07-AUG-2002  
DEFINITION Mus musculus, clone IMAGE:5008623, mRNA.  
ACCESSION BC026713  
VERSION BC026713.1 GI:19934327  
KEYWORDS HTC.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 5174)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (04-APR-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnarone, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,  
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 45 Row: b Column: 17  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis  
This clone has the following problem: no 5' EST match.  
FEATURES  
Location/Qualifiers  
1..5174  
/organism="Mus musculus"

```
/db_xref="taxon:10090"
/clone="IMAGE:5008623"
/tissue_type="Mammary tumor, MMTV-LTR/INT3 model, 5 month
old mouse, taken by biopsy."
/clone_lib="NCI_CGAP_Mam2"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
BASE COUNT      1356 a 1232 c 1246 g 1340 t
ORIGIN

Query Match      9.4%; Score 324.8; DB 11; Length 5174;
Best Local Similarity 58.0%; Pred. No. 8.3e-47;
Matches 612; Conservative 0; Mismatches 441; Indels 3; Gaps 2;

QY 1792 AACGAGTACTTCTAGCTTGCATTCAGGACATGATATGACCTTAAAGTGGAGTCCCG 1851
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1749 ACAAATATGTTTACATAGACCCGACGCAACTTCTTATGATCACAATATGGGATTC 1808
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1852 AGAGAGAACTTAGAGTTTGGGAAGTCTCTGGGCTTCGGGAGGTGATGAAC 1911
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1809 AGAACACGCTGAGTTTGGAAAGACATTTGGAGCTGGTGCCTTCGGGAAGTCTTG 1868
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1912 GCCACGGCTATGGCATTAGTAAACGGGAGTCTCAATTCAGGTGGCGTGAAGTCT 1971
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1869 GCCACTGCATATGGCTTGATTAACTCGGATGCTGCATGACAGTTGCCGTGAAGTCT 1928
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1972 AAAGAGAAGCTGACAGCTGTGAAGAGAGCTCTCATGTCGGAGCTCAAAATGATG 2031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1929 AAACCAAGTGCCATTAAACAGAAAGAGAGGCCCTAATGTCGGAATGAAGTCTG 1988
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2032 CACCTGGGACACATGACAACTGATCTGCTGGGGCATGCACACTGTCAGGGCCA 2091
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1989 TACCTGGGCAATCACATGAATATTGTGAACCTGTGGCGCATGCAGGTGGAGGGCC 2048
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2092 GTGTACTTGTATTTTCAATATTTGCTATGTGTGACCTCTCAACTACCTAAAGAA 2151
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2049 ACCCTGGTCATTACAGATATTTGCTATGTGTGATCTTTTGAATTTTGGAGGAAG 2108
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2152 AGAGAGAGTTTTCACAGGACATGGACAGAGATTTTAAAGACATATTTTCAGTCT 2211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2109 CGTACTCTGTTATTTCTCAAGCAAGAAAGACAGCGCAGAGCGGCACTTTATAAG 2168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2212 CTTACTTTCCAGGCACATTCAAATCCACGATCCCTGGTTTCAGAGAACTTCAGT 2271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2169 CTTCTGCACTCAACGGAGCTTCTCTGTGACAGTTCAAATGAATATATGACATG 2226
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2272 CGCGCCCTTGGATCAGCTCTCAGGGTTCAATGGGAATTCATTCATCTGAAGATGAG 2331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2227 CTGGCGTTTCTTACGTGGTCCCAACCAACAGACAGAGAGAGATCCCGCAAGATAG 2286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2332 GAATATGAACACAGAGAGGCTGGCAGAAAGAGAGAGAGATTTGAACGTCTGAGC 2391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2287 CGTACATAGAAAGAGACGTGACT-CCTGCCATCATGGAAGATGACAGCTGGCTCG 2345
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2392 TTTGAGACCTCTTGTCTTTCGTCACCAAGTGGCCAAAGGATGGAATTCCTGGAGT 2451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2346 CTGGATGATTGCTGAGCTTCTCTACCAAGTGGCCAAAGGATGCGGTCTCTCGCTCC 2405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2452 AAGTCGTGTGTCCACAGACCTGGCAGCCAGGAATGTGTGTGTCACCCACGGAAGTGT 2511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2406 AAGAATGTATTACAGAGATTTGGCAGCCAGGAATATCTCTCACTACGCGCGGATC 2465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2512 GTGAAGATCTGACTTTTGGATGGCCCGGAGACATCTCTGAGGACATCCAGCTACG 2571
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2466 ACAAGATTTGGCATTTTCGGGTAGCCAGAGACATCAGGAATGATTGCAATACGTGG 2525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2572 AGGGCAACGCACGCTCCGGTGAAGTGGAGTGGACCGACCGAGAGCTTATTTGAGG 2631
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2526 AAGGAAATGCACGACTGCCCTGAAAGTGGATGGCACAGAGAGCAATTTTCAGCTCG 2585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2632 TACACAATCAAGAGTGCACCTGCTCTACGGCATCTCTCTGGGAGATATTTTACACTG 2691
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

2586 TACACATTTGAAAGTGATCTCTGTCTTATGGGATTTTCTCTGGGAGCTTCTCTCTTA 2645
2692 GGTGTGAACCTTACCCTGGCATCTCTGTCTGACGCTAACTTCTATAACTGATTACAG 2751
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2646 GGAACGAGCCCTACCAGGATGCGGTGCGACTCCAGTCTTACAGATGATCAGGAA 2705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2752 GGATTTAAATGGAGCAGCCATCTTATGCCACAGAGGATATATCTTTGTAATGCAATCC 2811
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2706 GGTCTCCGATGTCAGCCCGGAGCAGCGCTGCCGAATGTATGACGCTCATGAAGACT 2765
2812 TCTCTGGCTTTTGACTCAAGGAGCGGCATCTTC 2847
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2766 TCTCTGGAGCTGACCCCTTGAAGAGGCCAACATTC 2801
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
BGI45808
LOCUS
DEFINITION
    uu98g07.yl Soares_mouse_NMGB_bcell Mus musculus cDNA clone
    IMAGE:3384757 5' similar to SW:FLT3_HUMAN P36888 FL CYTOKINE
    RECEPTOR PRECURSOR ; mRNA sequence.
BGI45808
ACCESSION
VERSION
KEYWORDS
SOURCE
    house mouse.
ORGANISM
    Mus musculus.
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
    1 (bases 1 to 394)
AUTHORS
    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
    National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
    Tumor Gene Index
JOURNAL
    Unpublished (1997)
COMMENT
    Contact: Robert Strausberg, Ph.D.
    Email: cgapbs-remail.nih.gov
    This clone is available royalty-free through LLNL; contact the
    IMAGE Consortium (info@image.llnl.gov) for further information.
    MGI:1089753
Trace considered overall poor quality
Seq primer: -40RP from Gibco
High quality sequence stop: 1.
FEATURES
    Location/Qualifiers
    1..394
    /organism="Mus musculus"
    /db_xref="taxon:10090"
    /clone_lib="Soares_mouse_NMGB_bcell"
    /lab_host="DH10B (phage-resistant)"
    /note="Organ: germinal B-cell; Vector: pT7T3D-Pac
    (Pharmacia) with a modified polylinker; Site_1: Not I;
    Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
    oligo(dT) primer [5'
    TGTTCACCAATCTGAAGTGGAGCGGCGCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
    T 3']; double-stranded cDNA was ligated to Eco RI
    adaptors (Pharmacia), digested with Not I and cloned into
    the Not I and Eco RI sites of the modified pT7T3 vector.
    Library is normalized; constructed by Bento Soares and
    M.Fatima Bonaldo."
BASE COUNT      95 a 86 c 109 g 104 t
ORIGIN

Query Match      9.1%; Score 314; DB 12; Length 394;
Best Local Similarity 94.2%; Pred. No. 1.7e-44;
Matches 326; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 2447 AGTTCAAGTCGTGTCTCCACAGAGACTGGCAGCCAGGAATGTGTGTGTGTGTGTGT 2506
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 48 AGTTAAGTCGCTGTCTCCACAGAGACTGGCAGCCATGATGTGTGTGTGTGTGTGTGTGT 107
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2507 AGTGTGTGAAGATCTCTGTGACTTTTGGACTGGCCGCGGACACATCTCTGAGCGACTCC 2566
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 108 AGTGTGTGAAGATCTCTGTGACTTTTGGACTGGCCCTAGACATGCTGAGCGACTTAC 167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```





Gencore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:18:14 ; Search time 60.9693 Seconds  
(without alignments)  
2168.053 Million cell updates/sec

Title: US-09-919-408-2

Perfect score: 5264

Sequence: 1 MRALAQRSDRRLLLVLSV.....RGGLRAQSPQKVIHRRS 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A\_Geneseq\_101002:\*
- 1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
  - 2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
  - 3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
  - 4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
  - 5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
  - 6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
  - 7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
  - 8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
  - 9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
  - 10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
  - 11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
  - 12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
  - 13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
  - 14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
  - 15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
  - 16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
  - 17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
  - 18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
  - 19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
  - 20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
  - 21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
  - 22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
  - 23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	ID	Description
1	5264	100.0	992	14 AAR37502	Murine flk-2. Mus
2	5264	100.0	992	16 AAR67815	Flk2 receptor prot
3	5264	100.0	992	16 AAR67535	Mouse flk-2. Mus
4	5264	100.0	992	17 AAR97418	Human foetal liver
5	5264	100.0	992	18 AAW19874	Murine flk-2 recep
6	5264	100.0	992	20 AAY08616	Murine flk-2 prote
7	5256	99.8	992	14 AAR44994	Murine flk-2 recep
8	5256	99.8	992	14 AAR31375	Murine flk-2. Mus
9	5239	99.5	992	13 AAR28038	Murine flk-2. Mus
10	5102	96.9	1000	16 AAR81868	Flk2/flt3 tyrosine

11	4436.5	84.3	993	16 AAR75961	Human STR-1. Homo
12	4429.5	84.1	993	16 AAR67816	Flk2 receptor prot
13	4429.5	84.1	993	16 AAR67536	Human flk-2. Homo
14	4429.5	84.1	993	17 AAR97419	Murine foetal live
15	4429.5	84.1	993	18 AAW19873	Human flk-2 recep
16	4429.5	84.1	993	20 AAY08617	Human flk-2 protei
17	4426.5	84.1	993	14 AAR37503	Human flk-2. Homo
18	4421.5	84.0	993	14 AAR44995	Human flk-2 recep
19	4420.5	84.0	993	16 AAR81869	Human Flk2/flt3 ty
20	4417.5	83.9	1167	14 AAR31376	Human flk-2. Homo
21	4409.5	83.8	983	19 AAW63588	Human receptor typ
22	4408	83.7	986	19 AAW63587	Human receptor typ
23	4408	83.7	986	19 AAW63589	Human receptor typ
24	4404	83.7	994	19 AAW63586	Human receptor typ
25	2806.5	53.3	665	19 AAW63585	Human receptor typ
26	2373.5	45.1	749	19 AAW78002	Protein PHON32390.
27	2226.5	42.3	481	15 AAR47579	Soluble Flk-2. Mu
28	1223.5	23.2	972	23 AAW11935	Colony stimulating
29	1223	23.2	975	22 AAE07144	Murine Kit/stem ce
30	1223	23.2	975	22 AAE07148	Mutant murine Kit/
31	1221.5	23.2	972	23 AAW79039	Human macrophage c
32	1221.5	23.2	972	23 AAW11936	Colony stimulating
33	1221.5	23.2	972	23 AAW11941	Colony stimulating
34	1221.5	23.2	972	23 AAW73585	Colony stimulating
35	1221.5	23.2	1055	22 ABG15479	Novel human diagno
36	1220.5	23.2	972	23 AAW11934	Colony stimulating
37	1220.5	23.2	972	23 AAW11937	Colony stimulating
38	1219.5	23.2	972	23 AAW11940	Colony stimulating
39	1216.5	23.1	972	23 AAW11938	Colony stimulating
40	1215	23.1	977	21 AAY51322	Bovine c-Kit bk-1
41	1212.5	23.0	972	23 AAW11939	Colony stimulating
42	1197.5	22.7	2129	22 ABG15478	Novel human diagno
43	1197.5	22.7	2129	22 ABG20749	Novel human diagno
44	1196	22.7	976	22 AAE07145	Human Kit/stem cel
45	1196	22.7	976	22 AAW00375	Human stem cell gr

ALIGNMENTS

RESULT 1  
AAR37502  
ID AAR37502 standard; Protein; 992 AA.  
XX AAR37502;  
XX  
XX 19-OCT-1993 (first entry)  
DT  
DE Murine flk-2.  
XX

XX Murine; receptor; protein; tyrosine kinase; pTK; flk-2; primitive;  
KW hematopoietic cell; mature; family; conserved; region;  
KW catalytic domain; c-kit; fetal liver kinase; flk; fetal; spleen;  
KW thymus; adult; brain; bone marrow; multipotential; CFU-Blast colony;  
KW hierarchy; transduction; T-lymphoid; lineage.  
XX  
XX Mus musculus.

XX	Key	Location/Qualifiers
XX	Peptide	1..27
FT	Protein	/note= "Hydrophobic leader sequence"
FT	Protein	/note= "Mature murine flk-2"
FT	Domain	28..544
FT	Region	/note= "Extracellular receptor domain"
FT	Region	545..564
FT	Domain	/note= "Transmembrane region"
FT	Domain	565..992
FT	Domain	/note= "Intracellular catalytic domain"
FT	Domain	618..623
FT	Domain	/note= "Catalytic sub-domain"
FT	Domain	811..819
FT	Domain	/note= "Catalytic sub-domain"

FT	Domain	832..834		301	MSYSTNRTMIRILLAFVSSVGRNDTGYITC	SSSKHPQSALVTILEKGFINATSSQEEY	360
FT	FT	/note= "Catalytic sub-domain"					
FT	Domain	857..862		361	EIDPYEKFCFSVREFKAYPRICTWIFSOAS	PCGEORGLDGYISIKSCDHNKPKGEYIFY	420
FT	FT	/note= "Catalytic sub-domain"					
FT	Domain	872..878		361	EIDPYEKFCFSVREFKAYPRICTWIFSOAS	PCGEORGLDGYISIKSCDHNKPKGEYIFY	420
FT	FT	/note= "Catalytic sub-domain"					
FT	Region	736..812		421	AENDDAQFTKMFNLIRKKPOVLANASASQ	SCSSDGYPLPSWTWKKSDKSPNCTEELP	480
FT	FT	/note= "flk-2 signature sequence"					
XX	PN	WO9310136-A.		421	AENDDAQFTKMFNLIRKKPOVLANASASQ	SCSSDGYPLPSWTWKKSDKSPNCTEELP	480
XX	XX						
XX	PD	27-MAY-1993.		481	EGVWNKKANRNVFGQWSSSTLNMSKAGLL	VKCCAYNSMGTSCETIFLNSPGPFPTQ	540
XX	XX						
XX	PF	16-NOV-1992;		481	EGVWNKKANRNVFGQWSSSTLNMSKAGLL	VKCCAYNSMGTSCETIFLNSPGPFPTQ	540
XX	XX						
XX	PR	15-NOV-1991;		541	DNISFYATIGLCLPFIVVLIVLICHKFK	KQFYESOLOMIQVTPGLDNEYFYVDFRDEY	600
XX	XX						
XX	PA	(UYPR-) UNIV PRINCETON.		541	DNISFYATIGLCLPFIVVLIVLICHKFK	KQFYESOLOMIQVTPGLDNEYFYVDFRDEY	600
XX	PI	Lemischka IR;					
XX	XX						
XX	XX	WPI: 1993-182479/22.		601	DLKWEFFPRENLEFGKVLGSGAFGRVMNA	TAYGISKTGYISIQVAVKMLKEKADSCKEALM	660
XX	DR	N-PSDB; AAQ40914.		601	DLKWEFFPRENLEFGKVLGSGAFGRVMNA	TAYGISKTGYISIQVAVKMLKEKADSCKEALM	660
XX	XX						
XX	PT	Totipotent haematopoietic stem cell receptors, their ligands and		661	SELKMMTHLGHHDNIVNLLGACTLGGPV	LYLIFEYCCYGDLLNLYLSKREKPHRTWTEIFK	720
XX	PT	DNA sequences - for treating anaemia(s) and bone marrow damage					
XX	PT	due to e.g. cancer chemotherapy or radiotherapy		661	SELKMMTHLGHHDNIVNLLGACTLGGPV	LYLIFEYCCYGDLLNLYLSKREKPHRTWTEIFK	720
XX	XX						
XX	PS	Claim 37; Fig 1a; 127pp; English.		721	EHNFSSYPTFOAHNSNMPGSGREVOLHP	LDQLSGFNGNSIHSEDEIEYENOKRLAESEE	780
XX	XX						
XX	CC	This sequence represents the murine receptor protein tyrosine kinase		721	EHNFSSYPTFOAHNSNMPGSGREVOLHP	LDQLSGFNGNSIHSEDEIEYENOKRLAESEE	780
XX	CC	(pTK), flk-2. The nucleic acid encoding this receptor is expressed					
XX	CC	in primitive hematopoietic cells and not in mature hematopoietic		781	EDLNVLPFEDLLCFAYQVAKGMEFLFKS	CVHRDLAARNLVTHGKVKVVICDFGLARDIL	840
XX	CC	cells. Members of this family of pTK's can be recognised by the		781	EDLNVLPFEDLLCFAYQVAKGMEFLFKS	CVHRDLAARNLVTHGKVKVVICDFGLARDIL	840
XX	CC	conserved amino acid regions in the catalytic domain. This family					
XX	CC	of pTK's also contains c-kit. These new receptors are termed fetal		841	SDSSYVVRGNARLPVKWMAPELSEFEG	IYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN	900
XX	CC	liver kinases (flk's) after the tissue in which they were discovered.		841	SDSSYVVRGNARLPVKWMAPELSEFEG	IYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN	900
XX	CC	flk-2 is also expressed in fetal spleen, fetal thymus, adult brain					
XX	CC	and adult bone marrow. flk-2 is expressed in individual multipotential		901	FYKLIQSGFKMEQPFYATGEGIVFMQSC	WAFDSRRKPSFPNLTSLGCOLAEAEACIRT	960
XX	CC	CCFU-Blast colonies capable of generating numerous multilineage colonies		901	FYKLIQSGFKMEQPFYATGEGIVFMQSC	WAFDSRRKPSFPNLTSLGCOLAEAEACIRT	960
XX	CC	upon replating. It is likely therefore, that flk-2 is expressed in					
XX	CC	the entire primitive portion of the hematopoietic hierarchy. This is		961	SIHLPKQAAPQORGLRAQSPQRQVKIHR	RS 992	
XX	CC	consistent with flk-2 being important in transducing putative self-					
XX	CC	renewal signals from the environment. flk-2 is the first receptor		961	SIHLPKQAAPQORGLRAQSPQRQVKIHR	RS 992	
XX	CC	pTK known to be expressed in the T-lymphoid lineage.					
XX	XX						
XX	Sequence	992 AA;					
XX	Query Match	100.0%; Score 5264; DB 14; Length 992;					
XX	Best Local Similarity	100.0%; Pred. NO. 0;					
XX	Matches 992; Conservative	0; Mismatches	0; Indels	0; Gaps	0;		
QY	1	MRALQSRDRLLLLVLSVMILETVTNQDLP	VTKCVLISHENNGSSAGKPSRYMRVRS	60			
DB	1	MRALQSRDRLLLLVLSVMILETVTNQDLP	VTKCVLISHENNGSSAGKPSRYMRVRS	60			
QY	61	PEDLOCTPRQSGTVEATVEAESGITLOV	QLATPGDLSCLWVFKHSSLGCQPHD	120			
DB	61	PEDLOCTPRQSGTVEATVEAESGITLOV	QLATPGDLSCLWVFKHSSLGCQPHD	120			
QY	121	LQNRGIVSMALNTVTEQAGELLLHIOSE	RANYTVLFTVNVNRTQLYVLRPRPYFRKMENQ	180			
DB	121	LQNRGIVSMALNTVTEQAGELLLHIOSE	RANYTVLFTVNVNRTQLYVLRPRPYFRKMENQ	180			
QY	181	DALLCISEGVPEPTVWVLCSSHRESCKE	GPVAVRKEKVLHELFGTDIRCCARNALGR	240			
DB	181	DALLCISEGVPEPTVWVLCSSHRESCKE	GPVAVRKEKVLHELFGTDIRCCARNALGR	240			
QY	241	ECTKLFTIDLNQAPQSTLPQLFLKVGEP	PLWIRCKAIHNVHGFGLTWELEDALEEGSYFE	300			
DB	241	ECTKLFTIDLNQAPQSTLPQLFLKVGEP	PLWIRCKAIHNVHGFGLTWELEDALEEGSYFE	300			
QY	301	MSYSTNRTMIRILLAFVSSVGRNDTGYIT	CSSSKHPQSALVTILEKGFINATSSQEEY	360			

RESULT 2  
AAR67815  
ID AAR67815 standard; Protein; 992 AA.  
XX  
AC AAR67815;  
XX  
DT 16-AUG-1995 (first entry)  
XX  
DE Flk2 receptor protein-tyrosine-kinase.  
XX  
KW Mouse Flk2; receptor protein-tyrosine-kinase; primitive  
hematopoietic cell; fetal liver kinase; diagnostic ligand  
isolation; bone marrow disease therapy.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..27  
FT /note= "signal peptide"  
FT Protein 28..992  
FT /note= "mature protein"  
FT Domain 28..544  
FT /note= "extracellular receptor domain"  
FT Domain 545..564  
FT /note= "transmembrane region"

FT Domain 565..992 /note= "intracellular catalytic domain" 360  
FT Domain 618..623 /note= "catalytic sub-domain" 420  
FT Domain 811..819 /note= "catalytic sub-domain" 420  
FT Domain 832..834 /note= "catalytic sub-domain" 480  
FT Domain 857..862 /note= "catalytic sub-domain" 480  
FT Domain 872..878 /note= "catalytic sub-domain" 540  
FT Region 709..785 /note= "signature sequence typical of Flk2" 540  
XX  
PN W09500554-A. 600  
XX  
XX 05-JAN-1995. 600  
XX 17-JUN-1994; 94WO-US06944. 660  
XX 18-JUN-1993; 93US-0080244. 660  
PR 21-JUN-1993; 93US-0081508. 720  
PR 23-NOV-1993; 93US-0157490. 720  
XX (UUPR-) UNIV PRINCETON. 720  
XX Lemischka IR; 780  
XX WPI; 1995-052014/07. 780  
DR N-PSDB; AAQ81012. 780  
XX  
XX Ligand for receptor protein tyrosine kinase - useful for the 840  
PT stimulation of primitive haematopoietic stem cells causing 840  
PT proliferation and/or differentiation 840  
XX Disclosure; Fig 1a; 131pp; English. 900  
XX  
XX The sequence corresponds to a mouse Flk2 (fetal liver kinase) 900  
CC receptor protein-tyrosine-kinase, which is expressed in primitive 900  
CC hematopoietic cells but not in mature hematopoietic cells. The 900  
CC protein is useful in isolation of receptor ligands, which have 900  
CC applications in diagnosis of bone marrow disorders and in 900  
CC stimulating proliferation and/or differentiation of primitive 900  
CC hematopoietic stem cells. 900  
XX  
SQ Sequence 992 AA; 992

Query Match 100.0%; Score 5264; DB 16; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALAQRSDRRLLLVLSVMIETVTNODLPVIRKCVLISHENNGSSAGKPSRYMRVRS 60  
Db 1 MRALAQRSDRRLLLVLSVMIETVTNODLPVIRKCVLISHENNGSSAGKPSRYMRVRS 60  
Qy 61 PEDLOCTPRRQSEGVYEAATVEVAESGSLTQVQLATPGDLSCLVFKHSLGQCPHD 120  
Db 61 PEDLOCTPRRQSEGVYEAATVEVAESGSLTQVQLATPGDLSCLVFKHSLGQCPHD 120  
Qy 121 LQNRGIVSMALNVTTQAGEYLLHQISERANYTVLFTVNVDRDTOLYLRRPFRKMNQ 180  
Db 121 LQNRGIVSMALNVTTQAGEYLLHQISERANYTVLFTVNVDRDTOLYLRRPFRKMNQ 180  
Qy 181 DALLCISGVPEPTVWVLCSSHRSCKREGPAVVRKEKVLHFGTDIRCCARNALGR 240  
Db 181 DALLCISGVPEPTVWVLCSSHRSCKREGPAVVRKEKVLHFGTDIRCCARNALGR 240  
Qy 241 ECTKLTIDLNOAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELEDKALEGSGYFE 300  
Db 241 ECTKLTIDLNOAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFLTWELEDKALEGSGYFE 300  
Qy 301 MSTYSTNTRMTIRILLAFVSSVGRNDTGYYTCSSSKHPSQSALVTILEKGINATSSQEEY 360

Db 301 MSTYSTNTRMTIRILLAFVSSVGRNDTGYYTCSSSKHPSQSALVTILEKGINATSSQEEY 360  
Qy 361 EIDPYEKFCSVRFKAYPRIRCTWIFSQASFPCEQRLGEGYISIKPCDHKNRPGEYIFY 420  
Db 361 EIDPYEKFCSVRFKAYPRIRCTWIFSQASFPCEQRLGEGYISIKPCDHKNRPGEYIFY 420  
Qy 421 AENDDAQFTKMTFLNIRKKRPOVLANASASQASCSGSDGYPLPSWTWKKCSKSPNCTEIP 480  
Db 421 AENDDAQFTKMTFLNIRKKRPOVLANASASQASCSGSDGYPLPSWTWKKCSKSPNCTEIP 480  
Qy 481 EGYWKKKANRKYFGQWVSSSTLNMSBAGKGLLVKCCAYNSMGTSCTEIFLNSPGPFPIQ 540  
Db 481 EGYWKKKANRKYFGQWVSSSTLNMSBAGKGLLVKCCAYNSMGTSCTEIFLNSPGPFPIQ 540  
Qy 541 DNISFYATIGLCLPFIIVLIVLICHYKKQFRYESQLOMTOVTGPDNEFYFVDFRDEY 600  
Db 541 DNISFYATIGLCLPFIIVLIVLICHYKKQFRYESQLOMTOVTGPDNEFYFVDFRDEY 600  
Qy 601 DLKWEPPRENLEFGKVLGSCAFGRVMNATAYGISKTGVSIOQAVKMLKERADSCKEALM 660  
Db 601 DLKWEPPRENLEFGKVLGSCAFGRVMNATAYGISKTGVSIOQAVKMLKERADSCKEALM 660  
Qy 661 SELKMMTHLGHHDNIYNLLGACTLSGVPVLIIFYCVCYGDLLNLYLSKREKFRHRTWTEIFK 720  
Db 661 SELKMMTHLGHHDNIYNLLGACTLSGVPVLIIFYCVCYGDLLNLYLSKREKFRHRTWTEIFK 720  
Qy 721 EHNFSYPTFOAHNSNMPGSRREVQLHPDQLSGFNGNSIHSEDEIEYENQKRLAEEEE 780  
Db 721 EHNFSYPTFOAHNSNMPGSRREVQLHPDQLSGFNGNSIHSEDEIEYENQKRLAEEEE 780  
Qy 781 EDLNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840  
Db 781 EDLNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840  
Qy 841 SDSSYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPPYGPVVDAN 900  
Db 841 SDSSYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPPYGPVVDAN 900  
Qy 901 FYKLIQSGFKMEOPFYATEGIYFVMQSWAFDSKRKPSFPNLTSLGCLAEAEACIRT 960  
Db 901 FYKLIQSGFKMEOPFYATEGIYFVMQSWAFDSKRKPSFPNLTSLGCLAEAEACIRT 960  
Qy 961 SIHLPKQAAPQQRGGLRAQSPQVQKIHRS 992  
Db 961 SIHLPKQAAPQQRGGLRAQSPQVQKIHRS 992

RESULT 3  
AAR67535  
ID AAR67535 standard; Protein; 992 AA.  
XX  
AC AAR67535;  
XX AC  
DT 04-JUL-1995 (first entry)  
XX  
DE Mouse flk-2.  
XX  
KW Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;  
KW hematopoiesis; stem cell.  
XX  
OS Mus sp.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..27  
FT /label= Sig\_peptide  
FT Domain 28..544  
FT /label= Extracellular\_receptor\_domain  
FT Region 545..564  
FT /label= Transmembrane\_region  
FT Domain 565..992  
FT /label= Intracellular\_catalytic\_domain  
XX



PR 28-JUN-1991; 91US-0728913.  
 PR 15-NOV-1991; 91US-0793065.  
 PR 24-DEC-1991; 91US-0813593.  
 PR 26-JUN-1992; 92US-0906397.  
 PR 12-NOV-1992; 92US-0975049.  
 PR 30-APR-1993; 93US-0055269.  
 PR 31-OCT-1994; 94US-0252517.  
 XX  
 PA (UYPR-) UNIV PRINCETON.  
 XX  
 XX Lemischka IR;  
 PI  
 XX  
 DR WPI; 1996-392678/39.  
 DR N-PSDB; AAT38733.  
 XX  
 DR  
 XX  
 PT Anti-fœtal liver kinase 2 (flk-2) antibodies - useful in assays,  
 PT for isolating haematopoietic stem cells expressing receptor and for  
 PT obtaining ligands  
 XX  
 PS Claim 1; Columns 27-34; 50pp; English.  
 XX  
 CC The present sequence is human foetal liver kinase 2 (flk-2), a  
 CC protein tyrosine kinase. Isolated antibodies, pref. monoclonal,  
 CC raised against the extracellular portion of flk-2 can be used to  
 CC assay for flk receptors on the surface of primitive haematopoietic  
 CC stem cells, and to isolate positive cells. The antibodies can also  
 CC be used as, or to obtain ligands, which stimulate the proliferation  
 CC and/or differentiation of stem cells. The ligands can be used, e.g.  
 CC for treating anaemia, or bone marrow damage resulting from cancer  
 CC chemotherapy, or radiation.  
 XX  
 SQ Sequence 992 AA;

Query Match 100.0%; Score 5264; DB 17; Length 992;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRLLLLVLSVMIETVTNODLPVKVLISHENNGSSAGKPSRYMRVRS 60  
 DB 1 MRALQSRDRLLLLVLSVMIETVTNODLPVKVLISHENNGSSAGKPSRYMRVRS 60  
 QY 61 PEDLQTPRQSEGTVEAATVEAESGITLQVQLATPGDLSCLVFKHSSLGCPHF 120  
 DB 61 PEDLQTPRQSEGTVEAATVEAESGITLQVQLATPGDLSCLVFKHSSLGCPHF 120  
 QY 121 LQNRGIVSMALNVTTQAGEYLLHQSERANYTVLFTVNRDTQLYVLRPFRKMNQ 180  
 DB 121 LQNRGIVSMALNVTTQAGEYLLHQSERANYTVLFTVNRDTQLYVLRPFRKMNQ 180  
 QY 181 DALLCISEGVPEPTVEMVLCSSHRESCKEKGPAVVRKEEKLVELFCTDIRCCARNALGR 240  
 DB 181 DALLCISEGVPEPTVEMVLCSSHRESCKEKGPAVVRKEEKLVELFCTDIRCCARNALGR 240  
 QY 241 ECTKLFTIDLNOAPOSTLPQLFLKVGEPWIRCKAIVHNHFGTLWELEDKALEESYFE 300  
 DB 241 ECTKLFTIDLNOAPOSTLPQLFLKVGEPWIRCKAIVHNHFGTLWELEDKALEESYFE 300  
 QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTSCSSKHPSQSALVTILEKGFINATSSOEY 360  
 DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTSCSSKHPSQSALVTILEKGFINATSSOEY 360  
 QY 361 EIDPYEKFCFSVRFKAYPIRCTWIFSOAFCEQRLGLEDGYISIKFCDHKNKPGYIIFY 420  
 DB 361 EIDPYEKFCFSVRFKAYPIRCTWIFSOAFCEQRLGLEDGYISIKFCDHKNKPGYIIFY 420  
 QY 421 AENDDAQFTKMTLNIRKPKQVLANASQASQSDGYPLPSWTWKCKSDKSPNCTEEIP 480  
 DB 421 AENDDAQFTKMTLNIRKPKQVLANASQASQSDGYPLPSWTWKCKSDKSPNCTEEIP 480  
 QY 481 EGVWKNKANRKYFGOWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFFIQ 540  
 DB 481 EGVWKNKANRKYFGOWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFFIQ 540

QY 541 DNISFYATIGLCLPFIIVLVILVILCHIKYKQFRYESQLQMIOVTGLDNEYFYVDFRDY 600  
 DB 541 DNISFYATIGLCLPFIIVLVILVILCHIKYKQFRYESQLQMIOVTGLDNEYFYVDFRDY 600  
 QY 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660  
 DB 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISKTVGSIQVAVKMLKEKADSCKEALM 660  
 QY 661 SELKMMTHLGHHDNIVNLLGACTILSGPVYLIFECYCCYGDLLNLYLSRKREKPHRTWTEIFK 720  
 DB 661 SELKMMTHLGHHDNIVNLLGACTILSGPVYLIFECYCCYGDLLNLYLSRKREKPHRTWTEIFK 720  
 QY 721 EHNFSYPTFOAHSSNMPGSRREVQLHPPDLQSLGFGNGNSTHSEDEIYEYENOKRLAEDEE 780  
 DB 721 EHNFSYPTFOAHSSNMPGSRREVQLHPPDLQSLGFGNGNSTHSEDEIYEYENOKRLAEDEE 780  
 QY 781 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGVKVKICDFGLARDIL 840  
 DB 781 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGVKVKICDFGLARDIL 840  
 QY 841 SDSYVVYVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNPPGIPVDAN 900  
 DB 841 SDSYVVYVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNPPGIPVDAN 900  
 QY 901 FYKLIOGFGKMEOPFVATEGIYFVMSQWAFDSRKRSPFNLTSLFLGCOLAEAEACIRT 960  
 DB 901 FYKLIOGFGKMEOPFVATEGIYFVMSQWAFDSRKRSPFNLTSLFLGCOLAEAEACIRT 960  
 QY 961 SIHLPKQAAPQORGGGLRAQSPQKVIHRERS 992  
 DB 961 SIHLPKQAAPQORGGGLRAQSPQKVIHRERS 992

RESULT 5  
 AAW19874  
 ID AAW19874 standard; Protein; 992 AA.  
 XX  
 AC AAW19874;  
 XX  
 DT 19-AUG-1997 (first entry)  
 XX  
 DE Murine flk-2 receptor.  
 XX  
 KW Human; fetal liver kinase 2; flk2; receptor protein tyrosine kinase;  
 KW ptk; liver; spleen; thymus; adult; brain; bone marrow; primitive portion;  
 KW haematopoietic hierarchy; extracellular domain; soluble form; ligand;  
 KW proliferation; differentiation; mammalian; haematopoietic stem cell;  
 KW macrocytic anaemia; aplastic anaemia; cancer; radiation; mouse.  
 XX  
 OS Mus musculus.  
 XX  
 FH Location/Qualifiers  
 FT Key  
 FT Peptide  
 FT Domain  
 FT Region  
 FT Domain  
 FT Domain  
 FT Active-site  
 FT Active-site  
 FT Active-site  
 FT Active-site  
 FT Active-site  
 FT Peptide

1..27  
 /note= "Signal peptide"  
 28..544  
 /note= "Extracellular receptor domain"  
 545..564  
 /note= "transmembrane region"  
 565..992  
 /note= "Intracellular catalytic domain"  
 618..623  
 /note= "Catalytic sub-domain"  
 811..819  
 /note= "Catalytic sub-domain"  
 832..834  
 /note= "Catalytic sub-domain"  
 857..862  
 /note= "Catalytic sub-domain"  
 872..878  
 /note= "Catalytic sub-domain"  
 709..785  
 /note= "Signature sequence characteristic of flk-2"





```
XX (UYPR-) UNIV PRINCETON.
XX Lemischka IR;
XX WPI; 1999-357194/30.
XX DR N-PSDB; AAX77514.
XX Isolating hematopoietic cells expressing fetal liver kinase 1
XX receptors
XX Disclosure; Fig 1a; 59pp; English.
XX This invention describes a novel method of isolating cells expressing
XX fetal liver kinase 1 (flk-1) receptors on their surface and comprises
XX binding the cells to a polyclonal or monoclonal antibody specific to
XX the flk-1 receptor and isolating the cells that have bound to the
XX any mammal but preferably a rat, mouse, rabbit or human. The proteins of
XX the invention belong to the receptor protein family. This sequence
XX represents the murine flk-2 protein which is used in the method of
XX the invention.
XX SQ Sequence 992 AA;
Query Match 100.0%; Score 5264; DB 20; Length 992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MRALAQRSDRRLLLVLSVMILETVTNQDLPVIRKCVLISHENNGSSAGKPSRYMRVRS 60
Db 1 MRALAQRSDRRLLLVLSVMILETVTNQDLPVIRKCVLISHENNGSSAGKPSRYMRVRS 60
Qy 61 PEDLOCTPRROSEGVYEAATVEAESSITLQVOLATPGDLSCLWVKHSSLGQCPHFD 120
Db 61 PEDLOCTPRROSEGVYEAATVEAESSITLQVOLATPGDLSCLWVKHSSLGQCPHFD 120
Qy 121 LONRGIVSMALINVTETQAGEVLLHQISERANYTVLFTVNVEDTOLYLRRPYFRKMNQ 180
Db 121 LONRGIVSMALINVTETQAGEVLLHQISERANYTVLFTVNVEDTOLYLRRPYFRKMNQ 180
Qy 181 DALLCISGEVPEPTVEWVLCSSHRESCKEEGPAVVRKEKVLHFGTDIRCCARNALGR 240
Db 181 DALLCISGEVPEPTVEWVLCSSHRESCKEEGPAVVRKEKVLHFGTDIRCCARNALGR 240
Qy 241 ECTKLTFTDLNOAQPOSTLPQLFLKVGEPILWIRCKAIHVNHGFLTWELDKALEEGSYFE 300
Db 241 ECTKLTFTDLNOAQPOSTLPQLFLKVGEPILWIRCKAIHVNHGFLTWELDKALEEGSYFE 300
Qy 301 MSTYSTNTMTIRILLAFVSSVGRNDTGYYTCSKKHPSQSALVTILEKGFINATSSQEEY 360
Db 301 MSTYSTNTMTIRILLAFVSSVGRNDTGYYTCSKKHPSQSALVTILEKGFINATSSQEEY 360
Qy 361 EIDPYEKFCFSVRKAYPRIRCTWIFSOASFCEORGLDGYISKFCDHKNKPGEYIF 420
Db 361 EIDPYEKFCFSVRKAYPRIRCTWIFSOASFCEORGLDGYISKFCDHKNKPGEYIF 420
Qy 421 AENDDAQFTKMTFLNIRKPKQVLANASQASCSDDGYPLPSWTWKCKSDKSPNCTEEIP 480
Db 421 AENDDAQFTKMTFLNIRKPKQVLANASQASCSDDGYPLPSWTWKCKSDKSPNCTEEIP 480
Qy 481 EGVWNNKANRKYFGQWVSSSTLNMSAGKGLLVKCCAYNSMGTSCETIFLNSPGFPFFIQ 540
Db 481 EGVWNNKANRKYFGQWVSSSTLNMSAGKGLLVKCCAYNSMGTSCETIFLNSPGFPFFIQ 540
Qy 541 DNISFYATIGLCPLFTVLVLICLCHKYKQFRYESQLOMIQVTPGLDNEYFYVDFRDY 600
Db 541 DNISFYATIGLCPLFTVLVLICLCHKYKQFRYESQLOMIQVTPGLDNEYFYVDFRDY 600
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVNMATAYGISKTGVSIOQAVKMLKEKADSCKEALM 660
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVNMATAYGISKTGVSIOQAVKMLKEKADSCKEALM 660
```

```
Qy 661 SELKMTHTLGHHDNIVNLIGACTLSGPVYLIFEYCCYGDLLNLRSKREKFRHTWTEIFK 720
Db 661 SELKMTHTLGHHDNIVNLIGACTLSGPVYLIFEYCCYGDLLNLRSKREKFRHTWTEIFK 720
Qy 721 EHNFSYPTFOAHSNNSMPGSRREVQLHPPDLQSLSGFNGNSIHSEDEIYEYENQKRLAEEEE 780
Db 721 EHNFSYPTFOAHSNNSMPGSRREVQLHPPDLQSLSGFNGNSIHSEDEIYEYENQKRLAEEEE 780
Qy 781 EDNLVLTFTEDLLCFAYQVAKGMELEFKSCVHRDLAARNVLTGHGVKVKICDFGLARDIL 840
Db 781 EDNLVLTFTEDLLCFAYQVAKGMELEFKSCVHRDLAARNVLTGHGVKVKICDFGLARDIL 840
Qy 841 SDSSYVVRGNARLPVKWMAPELSLFGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
Db 841 SDSSYVVRGNARLPVKWMAPELSLFGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
Qy 901 FYKLQSGFKMQPYATGEGYFVMSQWAFDSRKRPSFNLTSFLGCQLAEAEACIRT 960
Db 901 FYKLQSGFKMQPYATGEGYFVMSQWAFDSRKRPSFNLTSFLGCQLAEAEACIRT 960
Qy 961 SIHLPKQAAPQORGGLRAQSPORQVKIHRERS 992
Db 961 SIHLPKQAAPQORGGLRAQSPORQVKIHRERS 992
RESULT 7
AAR44994
ID AAR44994 standard; Protein; 992 AA.
XX
AC AAR44994;
XX
DT 27-JUN-1994 (first entry)
XX
DE Murine flk-2 receptor protein tyrosine kinase.
XX
KW Receptor protein tyrosine kinase; pTK family; foetal liver kinase;
KW primitive; totipotent; haematopoietic cell; stem cell; proliferation;
KW mfk; stromal cell.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FH Peptide 1..27
FT /label= signal_peptide
FT /note= "hydrophobic leader"
FT Protein 28..992
FT /label= flk-2
FT /note= "mature protein"
FT Domain 28..544
FT /label= extracellular_domain
FT Region 545..564
FT /label= transmembrane_region
FT Domain 565..992
FT /label= intracellular_catalytic_domain
FT Domain 618..623
FT /note= "catalytic subdomain"
FT Domain 811..819
FT /note= "catalytic subdomain"
FT Domain 832..834
FT /note= "catalytic subdomain"
FT Domain 857..862
FT /note= "catalytic subdomain"
FT Domain 872..878
FT /note= "catalytic subdomain"
XX
XX US5270458-A.
XX
XX 14-DEC-1993.
XX
XX 02-APR-1991; 91US-0679666.
XX
XX 02-APR-1991; 91US-0679666.
XX 28-JUN-1991; 91US-0728913.
```



XX 26-JUN-1992; 92WO-US05401.  
PF 28-JUN-1991; 91US-0728913.  
PR 15-NOV-1991; 91US-0793065.  
PR 24-DEC-1991; 91US-0813593.  
PR 02-APR-1992; 92WO-US02750.  
XX (UYPR-) UNIV PRINCETON.  
XX Lemischka IR;  
XX WPI; 1993-036323/04.  
DR N-PSDB; AAQ45247.  
XX Nucleic acid encoding receptor protein tyrosine kinase - allows  
PT development of ligands to stimulate proliferation and/or  
PT differentiation of mammalian haematopoietic stem cells  
XX Claim 5; Fig 1a; 78pp; English.  
XX This sequence represents a murine receptor protein tyrosine kinase  
CC which belongs to a new functional class of protein tyrosine kinases  
CC (PTKs). PTKs in this class are expressed in primitive mammalian  
CC hematopoietic (PNC) cells but not in mature hematopoietic cells (mHC).  
CC This protein is an example of a receptor PTK and is called fetal liver  
CC kinase 2 (flk-2). flk-2 is expressed in fetal liver, spleen and  
CC thymus, and adult brain and marrow. Expression of flk-2 mRNA occurs  
CC in the most primitive thymocyte subset, which is believed to be  
CC uncommitted. Therefore, thymocytes expressing flk-2 may be multi-  
CC potential. flk-2 is the first receptor tyrosine kinase known to be  
CC expressed in the T-lymphoid lineage.  
XX Sequence 992 AA;  
SQ

Query Match 99.8%; Score 5256; DB 14; Length 992;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 991; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MRALAQRSDRLLLLVLSMILETVTNODLPVICKVLISHENNGSSAGKPSYRMVRGS 60  
Db 1 MRALAQRSDRLLLLVLSMILETVTNODLPVICKVLISHENNGSSAGKPSYRMVRGS 60  
Qy 61 PEDLQCTPROSEGTVYEAATVEAESGSLTQVLATPGDLSCLWPKVHSHSGCQPHD 120  
Db 61 PEDLQCTPROSEGTVYEAATVEAESGSLTQVLATPGDLSCLWPKVHSHSGCQPHD 120  
Qy 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTVNVNVDTOYLVLRRPYFRKMNQ 180  
Db 121 LQNRGIVSMALNVTTQAGEYLLHIQSERANYTVLFTVNVNVDTOYLVLRRPYFRKMNQ 180  
Qy 181 DALLCISGEVPEPTVWVLCSSHRSCKEKGPAVVRKEKVLHFGTDIRCCARNALGR 240  
Db 181 DALLCISGEVPEPTVWVLCSSHRSCKEKGPAVVRKEKVLHFGTDIRCCARNALGR 240  
Qy 241 ECTKLFTDLNQAPOSTLPQLFKVGEPLWIRCKAIHVNHGGLTWELDKALEGSYFE 300  
Db 241 ECTKLFTDLNQAPOSTLPQLFKVGEPLWIRCKAIHVNHGGLTWELDKALEGSYFE 300  
Qy 301 MSTYSTNRTMIRILAFVSSVGRNDTGYTSSSKHPSQSALVTILEKGFINATSSQEEY 360  
Db 301 MSTYSTNRTMIRILAFVSSVGRNDTGYTSSSKHPSQSALVTILEKGFINATSSQEEY 360  
Qy 361 EIDPYEKFCFVRFKAYPRIRCTWIFSQAPCEQGLEGGYSISKFDHKNKPGYEIFY 420  
Db 361 EIDPYEKFCFVRFKAYPRIRCTWIFSQAPCEQGLEGGYSISKFDHKNKPGYEIFY 420  
Qy 421 AENDDAQTKMFTLIRKKPQVLANASQASCSGDPPLPSWTWKKCSKSPNCTEIP 480  
Db 421 AENDDAQTKMFTLIRKKPQVLANASQASCSGDPPLPSWTWKKCSKSPNCTEIP 480  
Qy 481 EGVNKKANRKFVFGOWSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540  
Db 481 EGVNKKANRKFVFGOWSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540

Db 481 EGVNKKANRKFVFGOWSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540  
Qy 541 DNISFYATIGLCLPFTVILVILICHKKYKQFRYESOLOMIQVTPGLDNEYFYVDFRDYEY 600  
Db 541 DNISFYATIGLCLPFTVILVILICHKKYKQFRYESOLOMIQVTPGLDNEYFYVDFRDYEY 600  
Qy 601 DLKWEFFPRENLEFGKVLGSGAFGRVNMATAYGISKTVGSIQAVKMLKADSCKEALM 660  
Db 601 DLKWEFFPRENLEFGKVLGSGAFGRVNMATAYGISKTVGSIQAVKMLKADSCKEALM 660  
Qy 661 SELKMMTHLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLNLYLRKREKFRHTWTEIFK 720  
Db 661 SELKMMTHLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLNLYLRKREKFRHTWTEIFK 720  
Qy 721 EHNFSYPTFOAHSNNSMPGSRVQLHPDQLDLSGNGNSIHSEDEIEYENOKRLAEEEE 780  
Db 721 EHNFSYPTFOAHSNNSMPGSRVQLHPDQLDLSGNGNSIHSEDEIEYENOKRLAEEEE 780  
Qy 781 EDNLNLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVVKICDFGLARDIL 840  
Db 781 EDNLNLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVVKICDFGLARDIL 840  
Qy 841 SDSSYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 900  
Db 841 SDSSYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 900  
Qy 901 FYKLIQSGFKMEQPPFATGEGIVFMQSWAFDSRKRPSPNLTSFLCQLAEAEACIRT 960  
Db 901 FYKLIQSGFKMEQPPFATGEGIVFMQSWAFDSRKRPSPNLTSFLCQLAEAEACIRT 960  
Qy 961 SIHLPKQAAPQORGGLRAOSPORQVAKIHRERS 992  
Db 961 SIHLPKQAAPQORGGLRAOSPORQVAKIHRERS 992

RESULT 9  
AAR28038  
ID AAR28038 standard; Protein; 992 AA.  
XX AAR28038;  
XX AC AAR28038;  
XX DT 15-MAR-1993 (first entry)  
XX DE Murine flk-2.  
XX KW Thymidine kinase; TK; haematopoietic; stem cells; proliferation;  
XX OS Mus musculus.  
XX Key Location/Qualifiers  
FH Peptide 1..27  
FT /note= "hydrophobic leader"  
FT Domain 28..544  
FT /note= "extracellular receptor domain"  
FT Region 545..564  
FT /note= "transmembrane region"  
FT Domain 565..992  
FT /note= "intracellular catalytic domain"  
FT Region 709..785  
XX /note= "signature sequence of flk-2"  
PN WO9217486-A.  
XX 15-OCT-1992.  
XX 02-APR-1992; 92WO-US02750.  
XX 02-APR-1991; 91US-0679666.  
XX 28-JUN-1991; 91US-0728913.  
XX 15-NOV-1991; 91US-0793065.  
XX 24-DEC-1991; 91US-0813593.



CC agonist antibodies able to bind to, and activate, flk2/flt3.  
CC Such antibodies can enhance repopulation of mature blood cell  
CC lineages following chemotherapy, radiotherapy, or bone marrow  
CC transplantation.  
XX  
SQ Sequence 1000 AA;

Query Match 96.9%; Score 5102; DB 16; Length 1000;  
Best Local Similarity 96.7%; Pred. No. 0;  
Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;

Qy 1 MRALAQRDRRLLLLVLSVMILETVTNODLPVICKVLISHENNGSSACKPSRYRMVRGS 60  
Db 1 MRALAQRDRRLLLLVLSVMILETVTNODLPVICKVLISHENNGSSACKPSRYRMVRGS 60  
Qy 61 PEDLOCTPRRSGEVYEAATVEVAESGITLQVOLATPGDLSCLVWFKHSSILGCPHFED 120  
Db 61 PEDLOCTPRRSGEVYEAATVEVAESGITLQVOLATPGDLSCLVWFKHSSILGCPHFED 120  
Qy 121 LQNRGIVSMALNVTETQAGEYLLHIQSERANYTVLFTVNVDRDTQLYLRLRPYFRKMENQ 180  
Db 121 LQNRGIVSMALNVTETQAGEYLLHIQSERANYTVLFTVNVDRDTQLYLRLRPYFRKMENQ 180  
Qy 181 DALLCISGEVPEPTVWVLCSSHRESCKEKGPAVVRKEKVLHELFGTDIRCCARNALGR 240  
Db 181 DALLCISGEVPEPTVWVLCSSHRESCKEKGPAVVRKEKVLHELFGTDIRCCARNALGR 240  
Qy 241 ECTKLFTIDLNOAQOSTLPOLFKVGEPLWIRCKAIHVNHGFLTWELEDKALEGSEYFE 300  
Db 241 ESTKLFTIDLNOAQOSTLPOLFKVGEPLWIRCKAIHVNHGFLTWELEDKALEGSEYFE 300  
Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYYTSCSSKHPQSOSALVTILEKGFINATSSQEEY 360  
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYYTSCSSKHPQSOSALVTILEKGFINATSSQEEY 360  
Qy 361 EIDPYEKFCSVRFKAYPRICTWIFSQAPECEQGLDGYSISKFCDHKNKPGYEIFY 420  
Db 361 EIDPYEKFCSVRFKAYPRICTWIFSQAPECEQGLDGYSISKFCDHKNKPGYEIFY 420  
Qy 421 AENDDAQTKFTLNIRKKPOVLANASQASCSDDGYPPLSWTWKCKSDKSPNCTEIP 480  
Db 421 AENDDAQTKFTLNIRKKPOVLANASQASCSDDGYPPLSWTWKCKSDKSPNCTEIP 480  
Qy 481 EGVNKKANRVFGOWSSSTLNSEAGKLLVKCCAYNSMGTSCETIFLNSPGPFPIQ 540  
Db 481 EGVNKKANRVFGOWSSSTLNSEAGKLLVKCCAYNSMGTSCETIFLNSPGPFPIQ 540  
Qy 541 DNISFYATIGLCLPFIVVLIVLICHKKQFRIESQLQMIQVTGPDNEIFYVDFRDEY 600  
Db 541 DNISFYATIGLCLPFIVVLIVLICHKKQFRIESQLQMIQVTGPDNEIFYVDFRDEY 600  
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVNNATAYGISTGYVSIQVAVMLKEKADSCKEALM 660  
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVNNATAYGISTGYVSIQVAVMLKEKADSCKEALM 660  
Qy 661 SELKMWTHLGHHDNVLNLLGACTLSGPVYLIFEYCCYGDLLNLRSKREKFRHTWTEIFK 720  
Db 661 SELKMWTHLGHHDNVLNLLGACTLSGPVYLIFEYCCYGDLLNLRSKREKFRHTWTEIFK 720  
Qy 721 EHNFSYPTFOAHNSMPGSRVQLHPPLDQLSGFNGSIHSEDEIEYENOKRLAEDEEE 780  
Db 721 EHNFSYPTFOAHNSMPGSRVQLHPPLDQLSGFNGSIHSEDEIEYENOKRLAEDEEE 780  
Qy 781 EDLNVLFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840  
Db 781 EDLNVLFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840  
Qy 841 SDSSTVVRGNARLPVKWMAPELSEGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900  
Db 841 SDSSTVVRGNARLPVKWMAPELSEGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900  
Qy 901 FYKLQSGFKMEQFPYATEGYFYVMQSWAFDSRRKPSFPNLTSLGCOLAEAEACIRT 960  
Db 901 FYKLQSGFKMEQFPYATEGYFYVMQSWAFDSRRKPSFPNLTSLGCOLAEAEACIRT 960

Db 901 FYKLQSGFKMEQFPYATEGYFYVMQSWAFDSRRKPSFPNLTSLGCOLAEAEAMYQN 960  
Qy 961 -----SIHLPKQAAPQORGLRAQSPQROVKIHRERS 992  
Db 961 MCGNVPEHPSIYONRRPLSREAGS-EPPSPQAQVKIHRERS 1000  
RESULT 11  
AAR75961  
ID AAR75961 standard; Protein; 993 AA.  
XX AAR75961;  
XX 29-DEC-1995 (first entry)  
XX Human STK-1.  
XX STK-1; receptor PTK; protein tyrosine kinase.  
XX Homo sapiens.  
FH Key Location/Qualifiers  
FT Peptide 1..23  
FT /label= signal  
FT /note= "unique PTK receptor sequence"  
FT Cleavage-site 25..26  
FT Domain 35..524  
FT /note= "part of extracellular domain contg.  
22 Cys residues marked on fig. 1"  
FT Modified-site 43..45  
FT /label= potential N-linked glycosylation site  
FT Modified-site 100..102  
FT /label= see above  
FT Modified-site 151..153  
FT /label= see above  
FT Modified-site 250..252  
FT /label= see above  
FT Modified-site 306..308  
FT /label= see above  
FT Modified-site 323..325  
FT /label= see above  
FT Modified-site 351..353  
FT /label= see above  
FT Modified-site 354..356  
FT /label= see above  
FT Modified-site 473..475  
FT /label= see above  
FT Modified-site 502..504  
FT /label= see above  
FT Region 542..562  
FT /label= transmembrane spanning region.  
FT Region 708..782  
FT /label= kinase insert region  
FT Domain 617..622  
FT /label= ATP binding domain  
FT Domain 835..840  
FT /label= WMAPES motifs  
FT /note= "cytoplasmic domain"  
FT Peptide 808..813  
FT /note= "used to design PCR oligos"  
FT Peptide 870..875  
FT /note= "used to design PCR oligos"  
XX W09519175-A.  
PN 20-JUL-1995.  
XX 06-JAN-1995; 95WO-US00176.  
XX 14-JAN-1994; 94US-0183211.  
XX (UJO ) UNIV JOHNS HOPKINS.  
XX (UYPE-) UNIV PENNSYLVANIA.  
PA

xx Civin C1, Gewirtz AM, Small D;  
 xx WPI; 1995-263709/34.  
 DR N-PSDB; AAQ91356.  
 xx Artificial STK-1 gene and gene-specific anti-sense oligo:nucleotide  
 PT used to treat neoplastic diseases and as bone marrow purging agents for  
 PT treating leukaemia and neoplasia  
 xx Disclosure; Fig 1; 66pp; English.  
 PS  
 xx The STK-1 gene encodes a receptor PTK which is expressed in  
 CC proliferating hematopoietic stem cells but not in quiescent stem  
 CC cells. The STK-1 gene is also expressed in certain malignant cells  
 CC of non-hematopoietic origin. An antisense oligo specific for STK-1  
 CC is an oligo having a sequence (i) capable of forming a stable  
 CC triplex with a portion of the STK-1 gene, or (ii) capable of forming  
 CC a stable duplex with a portion of an mRNA transcript of the STK-1  
 CC gene. Antisense oligos capable of forming a stable duplex with a  
 CC portion of a STK-1 mRNA transcript are given in AAQ91536 FT and in  
 CC AAQ91537 and AAQ91538. The antisense oligos of the invention are useful  
 CC in the treatment of hematologic malignancies characterized by  
 CC STK-1 expression. Several of the conserved domains of PTKs  
 CC including the ATP binding domain and the WMAPES motifs are  
 CC found in the STK-1 protein (see AAR75961 FT).  
 xx Sequence 993 AA;  
 SQ

Query Match 84.3%; Score 4436.5; DB 16; Length 993;  
 Best Local Similarity 84.2%; Pred. NO. 0;  
 Matches 837; Conservative 57; Mismatches 89; Indels 11; Gaps 4;

QY 1 MRALQSRDRLLLVLLVWLTETVNTQDLPVVKCVLISHENNGSSACKPSSYRVRGS 60  
 DB 1 MPALA-RDAGTVPLLVWTSAMTFTGNTQDLPVVKCVLHNKNDSSVGKSSYPVMS 59  
 QY 61 PEDLOCTPRQSEGVYEATVEAEGSITLQVQLATPGDLSCLVWFKHSLGCGPHFD 120  
 DB 60 PEDLGCALRQSSGVYEAAVEVDVSASITQLVLDAGNLSCLVWFKHSLNCPHFD 119  
 QY 121 LQNRGIVSMALNVTQTAGEYLLHIQSERANYTVLFTVNVRTQYLVLRPFRKMNQ 180  
 DB 120 LQNRGVSMVILKMTQAGEYLLFTQSEATNTILFTVSIRNTLLYTLRRPYFRKMNQ 179  
 QY 181 DALLCISEGVPEPTVWVLCSSHRECKEGPAPVVRKEKVLHELFGTDIRCCARNALGR 240  
 DB 180 DALVCISESVPEPIVWVLCDSQGESCKEESPAVVRKEKVLHELFGTDIRCCARNELGR 239  
 QY 241 ECTKLTIDLNOAPOSTLPOLFLKVGEPILWIRCKATHVNHGFLGTWELEDALEEGSYFE 300  
 DB 240 ECTRLFTIDLNOTPQTLPOLFLKVGEPILWIRCKATHVNHGFLGTWELEDALEEGSYFE 299  
 QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTTCSSSKHPSQSALVTILEKGFINATSSQEEY 360  
 DB 300 MSTYSTNRTMIRILLAFVSSVARNDTGYTTCSSSKHPSQSALVTIVEKGFINATNSSEY 359  
 QY 361 EIDPYEKFCSVRKAYPRICRTWISQSASFPCQRLGDEGYSISKPCDHKNRPGEYIF 420  
 DB 360 EIDPYEKFCSVRKAYPQICRTWISQSASFPCQRLGDEGYSISKPCDHKNRPGEYIF 419  
 QY 421 AENDDAQFTKMFPLNLRKRPVLANASASQASCSGYPPLSPWTKKCDKSPNCTEEIP 480  
 DB 420 AENDDAQFTKMFPLNLRKRPVLANASASQASCSGYPPLSPWTKKCDKSPNCTEEIT 479  
 QY 481 EGVNKKANKRVFGQWSSSTLNMSAGKLLVKCCAYNSMGTSCETIFLNSPGPFPIQ 540  
 DB 480 EGVNKKANKRVFGQWSSSTLNMSAIGKFLVKCCAYNSLGTSCETILLNSPGPFPIQ 539  
 QY 541 DNISFYATIGLCPLFFIVLVLVILCHYKKQFRYESQLOMIQVTPGLDNEYFYVDFRDEY 600  
 DB 540 DNISFYATIGVCLLFIWVLLTLLCHYKKQFRYESQLOMIQVTPGLDNEYFYVDFRDEY 599

QY 601 DLKWEFFPRENLEFGKVLGSGAFGRVNMATAYGISKTGVSIOQVAVKMLKEKADSCKEALM 660  
 DB 600 DLKWEFFPRENLEFGKVLGSGAFGRVNMATAYGISKTGVSIOQVAVKMLKEKADSSEREALM 659  
 QY 661 SELKMWTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYRSKREKFRHTWTEIFK 720  
 DB 660 SELKMWTLQSGSHENIVNLLGACTLSGPVYLIFEYCCYGDLLNLYRSKREKFRHTWTEIFK 719  
 QY 721 ENHFSYPTFOAHSSNMPGSRVQLHPDQSLGNGSNHSIHSEDEIYEYNOKRLAEDEE 780  
 DB 720 ENHFSYPTFOAHSSNMPGSRVQLHPDQSLGNGSNHSIHSEDEIYEYNOKRLAEDEE 777  
 QY 781 EDLNVLTFFEDLLCFAYQVAKGMEFEFKSCVHRDLAARNVLTHTGKVKICDFGLARDIL 840  
 DB 778 EDLNVLTFFEDLLCFAYQVAKGMEFEFKSCVHRDLAARNVLTHTGKVKICDFGLARDIM 837  
 QY 841 SDSYVVRGNARLPVKWMAPESEIFEGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 900  
 DB 838 SDSYVVRGNARLPVKWMAPESEIFEGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 897  
 QY 901 FYKLIQSGFKMEQPPFYATGEGYFVMQSWAFDSKRKPSFPNLTSLGCOLAEAEAC --- 957  
 DB 898 FYKLIQSGFKMEQPPFYATGEGYFVMQSWAFDSKRKPSFPNLTSLGCOLAEAEAMYQN 957  
 QY 958 ----IRTSIHLPKQAAPQORG-GLRAQSPQRQVK 986  
 DB 958 VDRVSECPHTYQNRPRPFSREMDLGLLSQAQVE 991

RESULT 12  
 AAR67816  
 ID AAR67816 standard; Protein; 993 AA.  
 AC AAR67816;  
 XX 18-AUG-1995 (first entry)  
 DE Flk2 receptor protein-tyrosine-kinase.  
 KW Human Flk2; receptor protein-tyrosine-kinase; primitive  
 KW hematopoietic cell; fetal liver kinase; diagnostic ligand  
 KW isolation; bone marrow disease therapy.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FT Peptide 1..27 /note= "signal peptide"  
 FT Protein 28..993 /note= "mature protein"  
 FT Domain 28..543 /note= "extracellular receptor domain"  
 FT Domain 544..563 /note= "transmembrane region"  
 FT Domain 564..993 /note= "intracellular catalytic domain"  
 XX W09500554-A.  
 XX 05-JAN-1995.  
 XX 17-JUN-1994; 94WO-US06944.  
 XX 18-JUN-1993; 93US-0080244.  
 PR 21-JUN-1993; 93US-0081508.  
 PR 23-NOV-1993; 93US-0157490.  
 XX (UYPR-) UNIV PRINCETON.  
 XX Lemischka IR;  
 XX WPI; 1995-052014/07.  
 DR N-PSDB; AAQ81013.

XX Ligand for receptor protein tyrosine kinase - useful for the  
PT stimulation of primitive haematopoietic stem cells causing  
PT proliferation and/or differentiation

XX Disclosure; Fig 1b; 131pp; English.

XX The sequence corresponds to a human Flk2 (fetal liver kinase)  
CC receptor protein-tyrosine-kinase, which is expressed in primitive  
CC hematopoietic cells but not in mature hematopoietic cells. The  
CC protein is useful in isolation of receptor ligands, which have  
CC applications in diagnosis of bone marrow disorders and in  
CC stimulating proliferation and/or differentiation of primitive  
CC hematopoietic stem cells.

XX Sequence 993 AA;

Query Match 84.1%; Score 4429.5; DB 16; Length 993;  
Best Local Similarity 84.1%; Pred. No. 0;  
Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

Qy 1 MRALAQRSDRLLLVLSVMIETVTNODLPVKVILSHENNGSSAGKPSYRWVRS 60  
Db 1 MPALA-RDAGTVPLLLVFSAMIFGTITNODLPVKVILSHENNGSSAGKPSYRWVRS 59  
Qy 61 PEDLOCTPRROSEGTVEAATVEAEGSITLOVOLATPGDLSCLVWFKHSLGCOPHFD 120  
Db 60 PEDLGCCALRPOSSGTVEAATVEAEGSITLOVOLATPGDLSCLVWFKHSLGCOPHFD 119  
Qy 121 LQNRGIVSMALNVTETQAGEYLLHIQSERANYTLFTVNVNRTOLYVLRPFYFKMENQ 180  
Db 120 LQNRGVSMVILKMTETQAGEYLLFIQSEATNYTLFTVNSIRNTLLYTLRRPFYFKMENQ 179  
Qy 181 DALICISEGPEPVEVWVLCSSHRESKEGPAVVRKEKVLHFLFGTDIRCCARNALGR 240  
Db 180 DALVCISESPEPVEVWVLCSSQGESKEGPAVVRKEKVLHFLFGTDIRCCARNELGR 239  
Qy 241 ECTKLFTIDLQAQPOSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWELEKALEGYSFE 300  
Db 240 ECTRLFTIDLNQTFQTLPLQLFKVGEPLWIRCKAVHVNHGFGLTWELEKALEGNYFE 299  
Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTCTSSSKHPSQSALVTILEKGFINATSSQBEY 360  
Db 300 MSTYSTNRTMIRILLAFVSSVARNDDGYTCTSSSKHPSQSALVTIVGKGFINATSSSEY 359  
Qy 361 EIDYEFECFSVRKAYPRIRCTWIFSOASPCBORGLEDGYSISKFDHKNKPGYIFY 420  
Db 360 EIDQYEEFCFSVRKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGYIFY 419  
Qy 421 AENDDAQFTKMTNIRKKPOVLANASASQSCSDGYPLPSWTWKKGSDKSPNCTEIP 480  
Db 420 AENDDAQFTKMTNIRKKPOVLAEASQSCSDGYPLPSWTWKKGSDKSPNCTEIT 479  
Qy 481 EGVNKKANRVFGOWVSSSTLNSEAGGLLVKCCAYNSMGTSCTETIFLNSPGFPFPIQ 540  
Db 480 EGVNKKANRVFGOWVSSSTLNSEALKGLVKKCAVNSLGTSETILLNSPGFPFPIQ 539  
Qy 541 DNISFYATIGLCLPIVVLVILCHKKYKQRYESQLOMIQVTPDNEIFYVDFRDYVEY 600  
Db 540 DNISFYATIGVCLLFIIVVTLILCHKKYKQRYESQLOMVQVTSQDNEIFYVDFREVEY 599  
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVNNATAYGISKTGYSIQVAVKMLKEKADSCKEALM 660  
Db 600 DLKWEFPRENLEFGKVLGSGAFGRVNNATAYGISKTGYSIQVAVKMLKEKADSSEREALM 659  
Qy 661 SELKMTLHGHNDNVNLLGACTLSGPVYLIFEYCCYGDLLNLYRSKREKPHRTWTEIFK 720  
Db 660 SELKMTQLGSHENIVNLLGACTLSGPVYLIFEYCCYGDLLNLYRSKREKPHRTWTEIFK 719  
Qy 721 EHNFSSTYPTQAHNSMPGSRVQLPPLDQLSGNSHSEDEIEYENOKRLAEDEE 780  
Db 720 EHNFSSTYPTQAHNSMPGSRVQLPPLDQLSGNSHSEDEIEYENOKRL - EEE 777

Qy 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840  
Db 778 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIM 837  
Qy 841 SDSYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPPYGPVVDAN 900  
Db 838 SDSYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPPYGPVVDAN 897  
Qy 901 FYKLIOGFKMQEPFVATEGIYFVWQSWAFDSRKRPSPNLTSFLGCCOLAEAEAC --- 957  
Db 898 FYKLIOGFKMQEPFVATEGIYIIMQSWAFDSRKRPSPNLTSFLGCCOLAEAEAWYQN 957  
Qy 958 ----IRTSIHLPKQAAPQORG-GLRAQSPQROVK 986  
Db 958 VDRVSECPHTYQNRPPFSREMDLGLLSPOAQVE 991

RESULT 13  
AAR67536  
ID AAR67536 standard; Protein; 993 AA.  
XX AAR67536;  
AC AAR67536;  
DT 04-JUL-1995 (first entry)  
XX Human flk-2.  
DE Human flk-2.  
XX Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;  
KW hematopoiesis; stem cell.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FT Peptide 1..27  
FT Domain 28..543 /label= Sig\_peptide  
FT Region 544..563 /label= Extracellular\_receptor\_domain  
FT Domain 564..993 /label= Transmembrane\_region  
FT Domain /label= Intracellular\_catalytic\_domain  
XX US5367057-A.  
XX 22-NOV-1994.  
XX 02-APR-1991; 91US-0679666.  
XX 02-APR-1991; 91US-0679666.  
XX 28-JUN-1991; 91US-0728913.  
XX 15-NOV-1991; 91US-0793065.  
XX 24-DEC-1991; 91US-0813593.  
XX 26-JUN-1992; 92US-0906397.  
XX 12-NOV-1992; 92US-0975049.  
XX 19-NOV-1992; 92US-0977451.  
XX 30-APR-1993; 93US-0055269.  
XX (UYPR-) UNIV PRINCETON.  
XX Lemischka IR;  
XX WPI; 1995-005894/01.  
XX N-PSDB; AAQ79069.  
XX Murine flk-2 receptor protein tyrosine kinase - used to stimulate  
PT proliferation and/or stimulation of primitive mammalian  
PT haematopoietic stem cells in vitro or in vivo.  
XX Disclosure; Fig. 2A-1F; 69pp; English.  
XX cDNAs encoding receptor protein tyrosine-kinases, mouse fetal liver  
CC kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAQ79068-70,  
CC respectively, and the deduced amino acid sequences in AAR67535-37,



CC respectively.

SQ Sequence 993 AA;

Query Match 84.1%; Score 4429.5; DB 16; Length 993;

Best Local Similarity 84.1%; Pred. No. 0;

Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

QY 1 MRALAQRSDRRLLLVLSVMILETVTNDLPVICKVLISHENNSSAGKPSRYMVRGS 60  
DB 1 MPALA-RDAGTVPLLVVFSAMIFGTTNDLPVICKVLINHNKNDSSVGKSSYPWVSES 59  
QY 61 PEDLOCTPRRQSEGTVEAATVEAERSGSITLQVQLATPEDLSCLWVFKHSSLCQPHFD 120  
DB 60 PEDGALCARPQSSGTVEAATAVEDVSASITLQVLDAPGNISCLWVFKHSSLCQPHFD 119  
QY 121 LQNRGIVSMAILNVTQAGEYLLHQTOSERANYTVLTVNVRTQYLVLRPPYFKRMENQ 180  
DB 120 LQNRGVSMVILKMTQAGEYLLFTQSEATNTYLTFSIRNTLLYTLRPPYFKRMENQ 179  
QY 181 DALLCISEGVPPEPTVEMVLCSSHRECKEKGPAVVRKEKVLHFGTDIRCCARNALGR 240  
DB 180 DALVCLISEVPEPIVWVLCDSGESCCKEESPAVVRKEKVLHFGTDIRCCARNELGR 239  
QY 241 ECTKLTIDLNAQOSTLPQLFLKVGELPWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300  
DB 240 ECTRLFTIDLNOTPQTLPLQLFLKVGELPWIRCKAVVNVHGFGLTWELENKALEEGNYFE 299  
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTCCSSSKHPSQSALVTILEKGFINATSSQEEY 360  
DB 300 MSTYSTNRTMIRILLAFVSSVGRNDGYTCCSSSKHPSQSALVTIVGKGFINATNSSEY 359  
QY 361 EIDPYEKFCSVRFKAYPRICRTWISQASFCPEORGLDGYISKFCDHKNKPGYIFY 420  
DB 360 EIDQYEEFCFSVRFKAYPQICRTWISQASFCPEORGLDGYISKFCNHKHPGXYIFH 419  
QY 421 AENDDAQFTKMTLNIRKRPQVLNANASQASCSGYPPLPSWTWKKCKSDKSPNCTEEIP 480  
DB 420 AENDDAQFTKMTLNIRKRPQVLNANASQASCSGYPPLPSWTWKKCKSDKSPNCTEEIT 479  
QY 481 EGVNKKANKRVFGQWVSSSTLNMSBAGKLLVKKCCAYNSMGTSCTETIFLNSPGPPFIQ 540  
DB 480 EGVNKKANKRVFGQWVSSSTLNMSBAGKLLVKKCCAYNSMGTSCTETIFLNSPGPPFIQ 539  
QY 541 DNISFATITGLCLPFIVLIVLICHYKQFRYESQIQMTOVTPGLDNEFYVDFRDEY 600  
DB 540 DNISFATITGVCLLFIWVLTLLCHYKQFRYESQIQMTOVTPGLDNEFYVDFRDEY 599  
QY 601 DLKWEPPRENLEFGKVLGSGAFGRVNVNATAYGISKTVSIOQVAVKMLKEKADSCKEALM 660  
DB 600 DLKWEPPRENLEFGKVLGSGAFGRVNVNATAYGISKTVSIOQVAVKMLKEKADSCSERALM 659  
QY 661 SELKMMTHLGHNDINVLNGLACTLSPGVLYIFCYCYGDLNLYLRSKREKFRHTWTEIFK 720  
DB 660 SELKMMTQLGSHENIVNLLGACTLSPGVLYIFCYCYGDLNLYLRSKREKFRHTWTEIFK 719  
QY 721 EHNFSYPTFOAHNSNMPGSRVQLHPDLQSLGNGNSIHSDEDEIYENQKRLAESEE 780  
DB 720 EHNFSYPTFOAHNSNMPGSRVQLHPDLQSLGNGNSIHSDEDEIYENQKRLAESEE 777  
QY 781 EDLNVLTFFEDLLCFAYQVAKGMFEFLFKSCVHRDLAARNVLTGHVKVVKICDFGLARDIL 840  
DB 778 EDLNVLTFFEDLLCFAYQVAKGMFEFLFKSCVHRDLAARNVLTGHVKVVKICDFGLARDIM 837  
QY 841 SDSYVVRGNARLPVKWAPESLFEIYTIKSDVSWYSGILLWEIFSGVNPYGPVDPAN 900  
DB 838 SDSYVVRGNARLPVKWAPESLFEIYTIKSDVSWYSGILLWEIFSGVNPYGPVDPAN 897  
QY 901 FYKLQSGKMEQOPFVATGIYVVMQSWAFDSKRKPSFNLTSFLGCQLAEAEAC--- 957  
DB 898 FYKLQSGKMEQOPFVATGIYVVMQSWAFDSKRKPSFNLTSFLGCQLAEAEAMQN 957  
QY 958 ----ITSTHLPKQAAPQORG-GLRAQSPORQVK 986

Db 958 VDGVRSECPHTYQNRPRFSPREMDLGLLSPOAQVE 991

RESULT 14

AAR97419

ID AAR97419 standard; Protein; 993 AA.

XX AAR97419;

XX 11-DEC-1996 (first entry)

XX Murine foetal liver kinase 2.

XX Murine; foetal liver kinase 2; flk-2; protein tyrosine kinase;

XX monoclonal; antibody; extracellular domain; receptor assay;

XX haematopoietic stem cell; ligand; stimulation; proliferation;

XX cancer chemotherapy; radiation.

XX Mus musculus.

XX Key Peptide Location/Qualifiers

XX Peptide 1..27 /label= sig\_peptide

XX Peptide 28..993 /label= mat\_peptide

XX Domain 28..544 /label= extracellular\_domain

XX Domain 545..564 /label= transmembrane\_domain

XX Domain 565..993 /label= intracellular\_domain

XX US5548065-A.

XX 20-AUG-1996.

XX 02-APR-1991; 91US-0679666.

XX 19-NOV-1992; 92US-0977451.

XX 02-APR-1991; 91US-0679666.

XX 28-JUN-1991; 91US-0728913.

XX 15-NOV-1991; 91US-0793065.

XX 24-DEC-1991; 91US-0813593.

XX 26-JUN-1992; 92US-0906397.

XX 12-NOV-1992; 92US-0975049.

XX 30-APR-1993; 93US-0055269.

XX 31-OCT-1994; 94US-0252517.

XX (UYPR-) UNIV PRINCETON.

XX Lemischka IR;

XX WPI; 1996-392678/39.

XX N-PSDB; AAT38734.

XX Anti-foetal liver kinase 2 (flk-2) antibodies - useful in assays.

XX for isolating haematopoietic stem cells expressing receptor and for

XX obtaining ligands

XX Claim 1; Columns 39-48; 50pp; English.

XX The present sequence is murine foetal liver kinase 2 (flk-2),

XX a protein tyrosine kinase. Isolated antibodies, pref. monoclonal,

XX raised against the extracellular portion of flk-2 can be used to

XX assay for flk receptors on the surface of primitive haematopoietic

XX stem cells, and to isolate positive cells. The antibodies can also

XX be used as, or to obtain ligands, which stimulate the proliferation

XX and/or differentiation of stem cells. The ligands can be used, e.g.

XX for treating anaemia, or bone marrow damage resulting from cancer

XX chemotherapy, or radiation.

SQ		Sequence	993 AA;
Query Match		84.1%;	Score 4429.5; DB 17; Length 993;
Best Local Similarity		84.1%;	Pred. No. 0;
Matches		836;	Conservative 57; Mismatches 90; Indels 11; Gaps 4;
Qy	1	MRALAQRSDRRLLLVLSVMIETVTNODLPVVKCVLISHENNGSSAGKPSRYMRVGRS	60
Db	1	MPALA-RDAGTVPLLVVFSAMIFCTITNODLPVVKCVLHKNNDSSVKGSSYPMVSES	59
Qy	61	PEDLOCTPQROSEGTVEATVEAEGSITLQVOLATPGDLSCLWPKHSLGCGPHFD	120
Db	60	PEDLGCLARPOSSGTVEAAAEVDVSASITLQVLDPAGNIGSLWPKHSLGCGPHFD	119
Qy	121	LONRGIVSMALNTVETOCAGEYLHIOGERANYTVLFTVNVVDLTQYVLRPYRKMENO	180
Db	120	LQNGVVSVMVLLKMTETOCAGEYLFIQSEATNYTILFTVSRNTLLYLRPYRKMENO	179
Qy	181	DALLCISGVPEPTVWVLCSSHRESCKEKPAAVVRKEEVKLVHELFGTDIRCCARNALGR	240
Db	180	DALVCISGVPEPIVWVLCDSGESCKEESPAAVVRKEEVKLVHELFGTDIRCCARNELGR	239
Qy	241	ECTKLFTIDLNOAQOSTLPQLFLKVGPELWIRCKAIHNVHGFGLTWELEDKALBEGSYFE	300
Db	240	ECTRLFTIDLNOTPTTLPLQLFLKVGPELWIRCKAVHNVHGFGLTWELENKALBEGNYFE	299
Qy	301	MSTYSTNRTMIRILLAFVSSVGRNDTGYTCTSSSKHPSQSALVTILEKGFINATSSQEEY	360
Db	300	MSTYSTNRTMIRILLAFVSSVARNDTGYTCTSSSKHPSQSALVTIVGKGFINATSSSEY	359
Qy	361	EIDPYEFCFSVREKAYPRICRTWIFSQAPEORGLDGYSTSKFCDHKNKPGYEIFY	420
Db	360	EIDQYEFECFSVREKAYPQIRCTWFSRKSPCEQKGLDNGYSTSKFCNHHKHQGEYIFH	419
Qy	421	AENDDAQFTKMTNIRKKPOVLANASQASCSGDGYPLPSWTWKKCSKSPNCTEIP	480
Db	420	AENDDAQFTKMTNIRKKPOVLAEASQASCSFSDGYPLPSWTWKKCSKSPNCTEIT	479
Qy	481	EGVNNKANRVFGQWSSSTLNSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPFPIQ	540
Db	480	EGVNNKANRVFGQWSSSTLNSEAIKGLVCCAYNSLGTSCETILLNSPGPFPIQ	539
Qy	541	DNISFYATIGLCPLFIVVLIIVLICHKKQFRIESQLOMIQVGTPLDNEYFYVDFRDYEE	600
Db	540	DNISFYATIGVCLLFIVVLTLLICHKKQFRIESQLOMQVGTGSSDNEYFYVDFREY	599
Qy	601	DLKWEFPRENLEFGKVLGSGAFGRVMNATAYIGSKTVGSIOVAVKMLKEKADSCKEALM	660
Db	600	DLKWEFPRENLEFGKVLGSGAFGRVMNATAYIGSKTVGSIOVAVKMLKEKADSSEREALM	659
Qy	661	SELKMTHLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLNLYRSKREKPHRTWTEIFK	720
Db	660	SELKMTQLGSHENIVNLGACTLSGPVYLIFEYCCYGDLLNLYRSKREKPHRTWTEIFK	719
Qy	721	EHNFSSYPTFOAHNSMPSGREVOLHPDLQSGFNGNSTHSEDEIEYENOKRLAESEE	780
Db	720	EHNFSSYPTFOAHNSMPSGREVQIHDPDSQISGLHNSFSEDEIEYENOKRL--EE	777
Qy	781	EDLNVLFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL	840
Db	778	EDLNVLFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIM	837
Qy	841	SDSYVVRGNARLPVKWMAPELSEGIYTIKSDVWSYGILLWEIFSLGVPNYPGIPVDAN	900
Db	838	SDSYVVRGNARLPVKWMAPELSEGIYTIKSDVWSYGILLWEIFSLGVPNYPGIPVDAN	897
Qy	901	FYKLIOGSKMEQEPYATEGIYFVWQSCWAFSDSRKRPFPNLTSLGCOLAEAEAC---	957
Db	898	FYKLIOGSKMEQEPYATEGIYIIMQSCWAFSDSRKRPFPNLTSLGCOLAEAEAMQYON	957
Qy	958	----IRTSIHLPKQAAPQORG--GLRAQSPQRQVK	986
Db	958	VDRGSECPHYQYONRPFPSREMDLGLLSPOAQVE	991

RESULT 15

AAW19873  
ID AAW19873 standard; Protein; 993 AA.

XX AAW19873;

XX 19-AUG-1997 (first entry)

XX Human flk-2 receptor.

Human; fetal liver kinase 2; flk2; receptor protein tyrosine kinase;  
PTK; liver; spleen; thymus; adult; brain; bone marrow; primitive portion;  
haematopoietic hierarchy; extracellular domain; soluble form; ligand;  
proliferation; differentiation; mammalian; haematopoietic stem cell;  
macrocytic anaemia; aplastic anaemia; cancer; radiation; mouse.

XX Homo sapiens.

Key Location/Qualifiers  
Peptide 1..27  
Protein /note= "Signal peptide"  
FT 28..993  
FT /note= "Mature flk-2"

XX US5621090-A.

XX 15-APR-1997.

XX 02-APR-1991; 91US-0679666.

XX 26-JUN-1992; 92US-0906397.

XX 02-APR-1991; 91US-0679666.

XX 28-JUN-1991; 91US-0728913.

XX 15-NOV-1991; 91US-0793065.

XX 24-DEC-1991; 91US-0813593.

XX (UYPR-) UNIV PRINCETON.

XX Lemischka IR;

XX WPI; 1997-235228/21.  
XX N-PSDB: AAT72117.

XX Protein containing the extracellular domain of human flk-2 - used  
for identification of primitive haematopoietic cell proliferation  
and differentiation stimulatory ligands, e.g. for treating anaemia

XX Claim 1; Fig 1B; 55pp; English.

XX This sequence represents human fetal liver kinase 2 (flk2). flk-2 is  
a receptor protein tyrosine kinase (PTK) and is important in transducing  
putative self-renewal signals from the environment. flk-2 is expressed  
in fetal liver, spleen, fetal thymus, adult brain and adult bone marrow,  
and it is thought that flk-2 is expressed in the entire primitive portion  
of the haematopoietic hierarchy. The invention concerns a recombinant  
nucleic acid, preferably mRNA, which encodes a protein containing only  
the extracellular domain of human flk-2 and lacking the flk-2 intra-  
cellular catalytic domain. The resultant protein represents a soluble  
form of flk-2 which is used to isolate specific ligands for flk-2. These  
ligands can be used to stimulate proliferation and/or differentiation of  
mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for  
treatment of macrocytic or aplastic anaemia or bone marrow damage caused  
by cancer treatment or radiation.

XX Sequence 993 AA;

Query Match 84.1%; Score 4429.5; DB 18; Length 993;  
Best Local Similarity 84.1%; Pred. No. 0;  
Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

Qy 1 MRALAQRSDRRLLLVLSVMIETVTNODLPVVKCVLISHENNGSSAGKPSRYMRVGRS 60



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:28:59 ; Search time 23.9879 Seconds  
(without alignments)  
1216.760 Million cell updates

Title: US-09-919-408-2

Perfect score:

Sequence: 1 MRALAQSRDRRLLLVLSV.....RGGLRAQSPQRQVKIHRERS 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

code processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

```
1: /cqn2_6/ptodata/1/iaa/5A_COMB.pcp:*
```

2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*

3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*

4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*

5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*

```
6: /cgn2_6/ptodata/1/iaa/backfiles1.pcp:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB	ID	Description
1	5264	100.0	992	1	US-07-813-593-2	Sequence 2, Appli
2	5264	100.0	992	1	US-07-977-451-2	Sequence 2, Appli
3	5264	100.0	992	1	US-07-946-507-2	Sequence 2, Appli
4	5264	100.0	992	1	US-08-252-517-2	Sequence 2, Appli
5	5264	100.0	992	1	US-07-906-397A-2	Sequence 2, Appli
6	5264	100.0	992	1	US-08-601-891-2	Sequence 2, Appli
7	5264	100.0	992	2	US-09-021-324-2	Sequence 2, Appli
8	5264	100.0	992	5	PCT-US92-02150-2	Sequence 2, Appli
9	5264	100.0	992	5	PCT-US92-05401-2	Sequence 2, Appli
10	5264	100.0	992	5	PCT-US92-09893-2	Sequence 2, Appli
11	5102	96.9	1000	1	US-08-222-299-2	Sequence 2, Appli
12	5102	96.9	1000	2	US-08-434-878-2	Sequence 2, Appli
13	5102	96.9	1000	5	PCT-US95-03718-2	Sequence 2, Appli
14	4436.5	84.3	993	1	US-08-222-299-4	Sequence 4, Appli
15	4436.5	84.3	993	2	US-08-434-878-4	Sequence 4, Appli
16	4436.5	84.3	993	5	PCT-US95-03718-4	Sequence 4, Appli
17	4430.5	84.2	1160	5	PCT-US92-05401-4	Sequence 4, Appli
18	4429.5	84.1	993	1	US-07-977-451-4	Sequence 4, Appli
19	4429.5	84.1	993	1	US-08-252-517-4	Sequence 4, Appli
20	4429.5	84.1	993	1	US-07-906-397A-4	Sequence 4, Appli
21	4429.5	84.1	993	1	US-08-601-891-4	Sequence 4, Appli
22	4429.5	84.1	993	2	US-09-021-324-4	Sequence 4, Appli
23	4429.5	84.1	993	5	PCT-US92-09893-4	Sequence 4, Appli
24	4408.5	83.7	993	1	US-08-183-211-2	Sequence 2, Appli
25	4408.5	83.7	993	5	PCT-US95-00176A-2	Sequence 2, Appli
26	2240.5	42.6	481	4	US-07-912-122-4	Sequence 4, Appli
27	2240.5	42.6	481	5	PCT-US93-06404-4	Sequence 4, Appli

	Best Local Similarity	100.0%;	Pred. No. 0;	Matches	992;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	MRLAQRSDRRLLLVLSVMILETTVTNODLPVKVLISHENNSSACKPSSYMRVGS	60										
DB	1	MRLAQRSDRRLLLVLSVMILETTVTNODLPVKVLISHENNSSACKPSSYMRVGS	60										
QY	61	PEDLQCTPRROSEGTVYAAATVEAAGSSITLQVQLATPGDLSCLWVFKHSLGCGPHD	120										
DB	61	PEDLQCTPRROSEGTVYAAATVEAAGSSITLQVQLATPGDLSCLWVFKHSLGCGPHD	120										
QY	121	LONRGIVSMATLNTETQAGVLLHIQSERANYTVLFTVNVRTDTOLYYLRRRYFRKMENQ	180										
DB	121	LONRGIVSMATLNTETQAGVLLHIQSERANYTVLFTVNVRTDTOLYYLRRRYFRKMENQ	180										
QY	181	DALLCISGEVPEPTVEWVLCSSHRESCKEEGPAVYKKEKVLHELFGTDIRCCARNALGR	240										
DB	181	DALLCISGEVPEPTVEWVLCSSHRESCKEEGPAVYKKEKVLHELFGTDIRCCARNALGR	240										
QY	241	ECTKLFTIDLNOAPOSTLQPLFLKVGEPILWIRCKAIHVNHGGLTWELDKALEEGSYFE	300										
DB	241	ECTKLFTIDLNOAPOSTLQPLFLKVGEPILWIRCKAIHVNHGGLTWELDKALEEGSYFE	300										
QY	301	MSTYSTNTMIRILLAFVSSVGRNDGYTCSSSKHPQSOSALVTILEKGFINATSSQEEY	360										
DB	301	MSTYSTNTMIRILLAFVSSVGRNDGYTCSSSKHPQSOSALVTILEKGFINATSSQEEY	360										
QY	361	EIDPYEKECFSVRFKAYPRICTWIFSOAFPCQORGLEDGYSISKFCDHKNKPGEYIFY	420										
DB	361	EIDPYEKECFSVRFKAYPRICTWIFSOAFPCQORGLEDGYSISKFCDHKNKPGEYIFY	420										
QY	421	AENDDAQFTKMTLNIIRKKPQVLANASQAQSSDGYPLPSWTWKKCSDKSPNCTEETIP	480										
DB	421	AENDDAQFTKMTLNIIRKKPQVLANASQAQSSDGYPLPSWTWKKCSDKSPNCTEETIP	480										
QY	481	EGVWNKKANRKFVGOWSSSTLNMSEAGLLVKCCAYNSMGTSCETIFLNSPGPPFFIQ	540										
DB	481	EGVWNKKANRKFVGOWSSSTLNMSEAGLLVKCCAYNSMGTSCETIFLNSPGPPFFIQ	540										
QY	541	DNISFYATIGLCPFTIVLVILICHYKKQFYESQLOMIQVTPGLDNEYFYVDFRDYIEY	600										
DB	541	DNISFYATIGLCPFTIVLVILICHYKKQFYESQLOMIQVTPGLDNEYFYVDFRDYIEY	600										
QY	601	DLKWEPFRENLEBFGKVLGSGAFGRVNMNATAYGISKTGVSIOQAVKMLKEKADSCKEALM	660										
DB	601	DLKWEPFRENLEBFGKVLGSGAFGRVNMNATAYGISKTGVSIOQAVKMLKEKADSCKEALM	660										
QY	661	SELKMTHTLGHNDINVLGACTLSGPVYLIFEYCCYGDLNLNLSRKREKHFHTWTETIFK	720										
DB	661	SELKMTHTLGHNDINVLGACTLSGPVYLIFEYCCYGDLNLNLSRKREKHFHTWTETIFK	720										
QY	721	ENHFSYPTFQAHSNNSMPGSRVQLHPPDLQDLSGFNGNSIHSEDEIYEYENQRLAESEE	780										
DB	721	ENHFSYPTFQAHSNNSMPGSRVQLHPPDLQDLSGFNGNSIHSEDEIYEYENQRLAESEE	780										
QY	781	EDLNVLTFEDLLCFAYQVAKGMEFLBFGKSCVHRDLAARNVLVTHGKVVKTCDFGLARDIL	840										
DB	781	EDLNVLTFEDLLCFAYQVAKGMEFLBFGKSCVHRDLAARNVLVTHGKVVKTCDFGLARDIL	840										
QY	841	SDSSYVVRGNARLPVKWMAPESLFEGYITTKSDVWSGILLWEIFSLGVNYPGIPVDAN	900										
DB	841	SDSSYVVRGNARLPVKWMAPESLFEGYITTKSDVWSG											

## RESULT 2

```
QY 61 PEDLOCTPRROSEGTVEAATVEAEGSITLQVOLATPGDLSCLWVFKHSSILGCOPHD 120
DB 61 PEDLOCTPRROSEGTVEAATVEAEGSITLQVOLATPGDLSCLWVFKHSSILGCOPHD 120
QY 121 LONRGIVSMALINVTETQAGEYLLHIQSERANYTLFTVNVNVRTQLYVLRFRYFRKMEQ 180
DB 121 LONRGIVSMALINVTETQAGEYLLHIQSERANYTLFTVNVNVRTQLYVLRFRYFRKMEQ 180
QY 181 DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEVLHELFGTDIRCCARNALGR 240
DB 181 DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEVLHELFGTDIRCCARNALGR 240
QY 241 ECTKLFTIDLNOAPOSTLPQLFLKVGEPWIRCKAITHVNHGFGLTWELEDKALEEGSYFE 300
DB 241 ECTKLFTIDLNOAPOSTLPQLFLKVGEPWIRCKAITHVNHGFGLTWELEDKALEEGSYFE 300
QY 301 MSTYSTNTRMTIRILLAFVSSVGRNDTGYTTCSSSKHPQSALVTILEKGFINATSSOEY 360
DB 301 MSTYSTNTRMTIRILLAFVSSVGRNDTGYTTCSSSKHPQSALVTILEKGFINATSSOEY 360
QY 361 BIDYEKFCFSVRKAYPRIRCTWIFSOASPCBOGLEDDGYSISKFCDHKNKPGYEIFY 420
DB 361 BIDYEKFCFSVRKAYPRIRCTWIFSOASPCBOGLEDDGYSISKFCDHKNKPGYEIFY 420
QY 421 AENDDAQFTKFTLNIRKKPOVLANASQASCSGSDGYPLPSWTWKKCDKSPNCTEETP 480
DB 421 AENDDAQFTKFTLNIRKKPOVLANASQASCSGSDGYPLPSWTWKKCDKSPNCTEETP 480
QY 481 EGVNKKANRRVFGOWVSSSTLNSEAGKGLVKKCAYNMGTSCTETIFLNSPGPPFFIQ 540
DB 481 EGVNKKANRRVFGOWVSSSTLNSEAGKGLVKKCAYNMGTSCTETIFLNSPGPPFFIQ 540
QY 541 DNISFYATIGLCLPIVIVLVILCHIKKQRYESQLOMIQVTPLDNEYFYVDFRDEY 600
DB 541 DNISFYATIGLCLPIVIVLVILCHIKKQRYESQLOMIQVTPLDNEYFYVDFRDEY 600
QY 601 DLKWEFFRENLEFGKVLGSGAGFRVNNATAGISKTGYSIQVAVKMLKEKADSCKEALM 660
DB 601 DLKWEFFRENLEFGKVLGSGAGFRVNNATAGISKTGYSIQVAVKMLKEKADSCKEALM 660
QY 661 SELKMTHLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLNLYRSKREKFRHTWTEIFK 720
DB 661 SELKMTHLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLNLYRSKREKFRHTWTEIFK 720
QY 721 EHNFSYPTFOAHSSNMPGSGREVOLHPPLDOLSGFNGNSIHSEDEIYENOKRLAESEE 780
DB 721 EHNFSYPTFOAHSSNMPGSGREVOLHPPLDOLSGFNGNSIHSEDEIYENOKRLAESEE 780
QY 781 EDLNVLTPEDLCLCFAYQVAKGMEFLFSCVHRDLAARNVLTGKVKYKICDFGLARDIL 840
DB 781 EDLNVLTPEDLCLCFAYQVAKGMEFLFSCVHRDLAARNVLTGKVKYKICDFGLARDIL 840
QY 841 SDSSVYVRGNARLPVKMMAPESLFGIYTIKSDVNSYIGILLWEIFSLGVNYPGIPVDAN 900
DB 841 SDSSVYVRGNARLPVKMMAPESLFGIYTIKSDVNSYIGILLWEIFSLGVNYPGIPVDAN 900
QY 901 FYKLQSGFKMEQPYATEGIFYVMQSCWAFDSRRKPSFPNLTSLGCOLAPAEACIRT 960
DB 901 FYKLQSGFKMEQPYATEGIFYVMQSCWAFDSRRKPSFPNLTSLGCOLAPAEACIRT 960
QY 961 SIHLPKQAAPQOGRGLRAQSPQOVKIHRS 992
DB 961 SIHLPKQAAPQOGRGLRAQSPQOVKIHRS 992
```

## RESULT 3

```
US-07-946-507-2
; Sequence 2, Application US/07946507
; Patent No. 5283354
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
```

```
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,507
; FILING DATE: 19920917
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEN-3-PPP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-946-507-2
```

```
Query Match 100.0%; Score 5264; DB 1; Length 992;
Best Local Similarity 100.0%; Pred. No.: 0;
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MRALAQRSDRRLLLVLSVMILETVTNODLPVVKCVLISHENNGSSAGKPPSSYRMVRS 60
DB 1 MRALAQRSDRRLLLVLSVMILETVTNODLPVVKCVLISHENNGSSAGKPPSSYRMVRS 60
QY 61 PEDLOCTPRROSEGTVEAATVEAEGSITLQVOLATPGDLSCLWVFKHSSILGCOPHD 120
DB 61 PEDLOCTPRROSEGTVEAATVEAEGSITLQVOLATPGDLSCLWVFKHSSILGCOPHD 120
QY 121 LONRGIVSMALINVTETQAGEYLLHIQSERANYTLFTVNVNVRTQLYVLRFRYFRKMEQ 180
DB 121 LONRGIVSMALINVTETQAGEYLLHIQSERANYTLFTVNVNVRTQLYVLRFRYFRKMEQ 180
QY 181 DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEVLHELFGTDIRCCARNALGR 240
DB 181 DALLCISEGVPEPTVEWVLCSSHRESCKEKGPAVVRKEEVLHELFGTDIRCCARNALGR 240
QY 241 ECTKLFTIDLNOAPOSTLPQLFLKVGEPWIRCKAITHVNHGFGLTWELEDKALEEGSYFE 300
DB 241 ECTKLFTIDLNOAPOSTLPQLFLKVGEPWIRCKAITHVNHGFGLTWELEDKALEEGSYFE 300
QY 301 MSTYSTNTRMTIRILLAFVSSVGRNDTGYTTCSSSKHPQSALVTILEKGFINATSSOEY 360
DB 301 MSTYSTNTRMTIRILLAFVSSVGRNDTGYTTCSSSKHPQSALVTILEKGFINATSSOEY 360
```

Qy 361 EIDPYEKFCFSVRKAYPRICRTWIFSQASFPCEQRLGEGYSISKFCDHKNKPGEIFY 420  
Db 361 EIDPYEKFCFSVRKAYPRICRTWIFSQASFPCEQRLGEGYSISKFCDHKNKPGEIFY 420  
Qy 421 AENDDAQFTKMTNIRKKPOVLANASQASCSGSDGYPLPSWTWKKCDKSPNCTEIP 480  
Db 421 AENDDAQFTKMTNIRKKPOVLANASQASCSGSDGYPLPSWTWKKCDKSPNCTEIP 480  
Qy 481 EGVNKKANRVFGQWSSVTLNNSKAGLLVKCCAYNSMGTSCTIFLNSPGFPFIQ 540  
Db 481 EGVNKKANRVFGQWSSVTLNNSKAGLLVKCCAYNSMGTSCTIFLNSPGFPFIQ 540  
Qy 541 DNISFYATIGLCLPFIVVLIVLICHKKQFRYESQLOMIQVTPGLDNEYFYVDFRDY 600  
Db 541 DNISFYATIGLCLPFIVVLIVLICHKKQFRYESQLOMIQVTPGLDNEYFYVDFRDY 600  
Qy 601 DLKWEFPRENLEFGKVLGSGAGFRVMNATAYGISTGYVSIQVAVKMLKADSCKEALM 660  
Db 601 DLKWEFPRENLEFGKVLGSGAGFRVMNATAYGISTGYVSIQVAVKMLKADSCKEALM 660  
Qy 661 SELKMWTHLGHHDNIVNLGACTLGGPVYLIFEYCCYGDLLNLYLSKREKFRHTWTEIFK 720  
Db 661 SELKMWTHLGHHDNIVNLGACTLGGPVYLIFEYCCYGDLLNLYLSKREKFRHTWTEIFK 720  
Qy 721 EHNFSYPTFOAHNSNMPGSRVQLHPPLDOLSGFNHNSIHSDEIEIYENOKRLAE 780  
Db 721 EHNFSYPTFOAHNSNMPGSRVQLHPPLDOLSGFNHNSIHSDEIEIYENOKRLAE 780  
Qy 781 EDLNVLFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGKVKIKCDFGLARDIL 840  
Db 781 EDLNVLFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGKVKIKCDFGLARDIL 840  
Qy 841 SDSSVVRGNARLPVKWMAPELFGIYTIKSDVNSYIGILLWEIFSLGVNYPGIPVDAN 900  
Db 841 SDSSVVRGNARLPVKWMAPELFGIYTIKSDVNSYIGILLWEIFSLGVNYPGIPVDAN 900  
Qy 901 FYKLIOGFKMEQPPYATGEGYFVMOQWAFDSRKRFPNLTSLGCOLAEAEACINT 960  
Db 901 FYKLIOGFKMEQPPYATGEGYFVMOQWAFDSRKRFPNLTSLGCOLAEAEACINT 960  
Qy 961 SIHLPKQAPOQGRGLRAQSPQOVKIHRS 992  
Db 961 SIHLPKQAPOQGRGLRAQSPQOVKIHRS 992

RESULT 4

US-08-252-517-2  
Sequence 2, Application US/08252517  
Patent No. 5548065  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Imclone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/252,517  
FILING DATE: 31-OCT-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/977,451  
FILING DATE: 19-NOV-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 992 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-252-517-2  
Query Match 100.0%; Score 5264; DB 1; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MRALAQRSDRRLLLVLSVMILEVTNODLPVICKVLISHENNSSAGKPSYRMVRS 60  
Db 1 MRALAQRSDRRLLLVLSVMILEVTNODLPVICKVLISHENNSSAGKPSYRMVRS 60  
Qy 61 PEDLOCTPRQSGTVYEATVEAESGSITLQVQLATPGDLSCLVWFKHSSLCQPHED 120  
Db 61 PEDLOCTPRQSGTVYEATVEAESGSITLQVQLATPGDLSCLVWFKHSSLCQPHED 120  
Qy 121 LQNRGTVSMALNVTTQAGEYLLHQSERANTVLTFTVNVRTQLYVLRPFRKMNQ 180  
Db 121 LQNRGTVSMALNVTTQAGEYLLHQSERANTVLTFTVNVRTQLYVLRPFRKMNQ 180  
Qy 181 DALLCISEGVPEPTVEWVLCSSHRESCKEKPAPVVRKEKVLHELFTDTRCCARNALGR 240  
Db 181 DALLCISEGVPEPTVEWVLCSSHRESCKEKPAPVVRKEKVLHELFTDTRCCARNALGR 240  
Qy 241 ECTKLTIDLNOAQOSTLPOLFLKVGEPMLIRCKA IHVNHGFGLTWELEKALEEGSYFE 300  
Db 241 ECTKLTIDLNOAQOSTLPOLFLKVGEPMLIRCKA IHVNHGFGLTWELEKALEEGSYFE 300  
Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYYTCSSSKHPSQSALVTILEKGFINATSSOEY 360  
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYYTCSSSKHPSQSALVTILEKGFINATSSOEY 360  
Qy 361 EIDPYEKFCFSVRKAYPRICRTWIFSQASFPCEQRLGEGYSISKFCDHKNKPGEIFY 420  
Db 361 EIDPYEKFCFSVRKAYPRICRTWIFSQASFPCEQRLGEGYSISKFCDHKNKPGEIFY 420  
Qy 421 AENDDAQFTKMTNIRKKPOVLANASQASCSGSDGYPLPSWTWKKCDKSPNCTEIP 480  
Db 421 AENDDAQFTKMTNIRKKPOVLANASQASCSGSDGYPLPSWTWKKCDKSPNCTEIP 480



Db 421 AENDDAQFTKMTLNIRKKPQVLANASQASCSDDGYPLPSWTWKCKSDKSPNCTEIP 480  
Qy 481 EGVNKKANKRVFGQWSSSTLNMSEAGKLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540  
Db 481 EGVNKKANKRVFGQWSSSTLNMSEAGKLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540  
Qy 541 DNISFYATIGLCLPIVVLVILVILCHYKQFRYESQLOMIQVTPGLDNEYFYVDFRDY 600  
Db 541 DNISFYATIGLCLPIVVLVILVILCHYKQFRYESQLOMIQVTPGLDNEYFYVDFRDY 600  
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOQVAVKMLKEKADSCKEALM 660  
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOQVAVKMLKEKADSCKEALM 660  
Qy 661 SELKMTHTLGHHDNIVNLGACTLSGVPYLIFFEYCCYGDLLNLYLSKREKPHRTWTEIFK 720  
Db 661 SELKMTHTLGHHDNIVNLGACTLSGVPYLIFFEYCCYGDLLNLYLSKREKPHRTWTEIFK 720  
Qy 721 EHNFSYPTFOAHNSMMPGSRREVQLHPDLQSLGNGNSIHSEDEIYEYENOKRLAE 780  
Db 721 EHNFSYPTFOAHNSMMPGSRREVQLHPDLQSLGNGNSIHSEDEIYEYENOKRLAE 780  
Qy 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840  
Db 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840  
Qy 841 SDSSVYVRGNARLPVKWAPESLFEIGIYTIKSDVWSYGILLWEIFSLGWNYPGIPVDAN 900  
Db 841 SDSSVYVRGNARLPVKWAPESLFEIGIYTIKSDVWSYGILLWEIFSLGWNYPGIPVDAN 900  
Qy 901 FYKLIOGFKMEQPYATEGYIFVMOQSWAFDSRRKPSFPNLTSLFGCOLAEAEACIRT 960  
Db 901 FYKLIOGFKMEQPYATEGYIFVMOQSWAFDSRRKPSFPNLTSLFGCOLAEAEACIRT 960  
Qy 961 SIHLPKQAAPQOORGLRAQSPORQVKIHRERS 992  
Db 961 SIHLPKQAAPQOORGLRAQSPORQVKIHRERS 992

## RESULT 5

US-07-906-397A-2  
; Sequence 2, Application US/07906397A  
; Patent No. 5621090  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCLONE SYSTEMS INCORPORATED  
; STREET: 180 VARICK STREET  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/906.397A  
; FILING DATE: 19920626  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813.593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793.065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728.913  
; FILING DATE: 28-JUN-1991

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/679.666  
; FILING DATE: 02-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28.601  
; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405  
; TELEFAX: 212-645-2054  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 992 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-906-397A-2

Query Match 100.0%; Score 5264; DB 1; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MRALAQSRDRRLLLVLSVMILETVTNQDLPVICKVLISHENNGSSAGKPSYRMVGRS 60  
Db 1 MRALAQSRDRRLLLVLSVMILETVTNQDLPVICKVLISHENNGSSAGKPSYRMVGRS 60  
Qy 61 PEDLOCTPRQSEGTVYEATVEAESGSITLQVQLATPGDLSCLMVFKHSSLGCCQPHFD 120  
Db 61 PEDLOCTPRQSEGTVYEATVEAESGSITLQVQLATPGDLSCLMVFKHSSLGCCQPHFD 120  
Qy 121 LQNRGIVSWAILNVTQAGEYLLHIQSEBRANVTVLFTVNVNRTQVLYLRRPFRKMEQ 180  
Db 121 LQNRGIVSWAILNVTQAGEYLLHIQSEBRANVTVLFTVNVNRTQVLYLRRPFRKMEQ 180  
Qy 181 DALLCISEGVPEPTVEWLVLCSSHRESCKEEGPAVVRKEEVLHEFGTDIRCCARNALGR 240  
Db 181 DALLCISEGVPEPTVEWLVLCSSHRESCKEEGPAVVRKEEVLHEFGTDIRCCARNALGR 240  
Qy 241 ECTKLTIDLNOAPQSTLPQLFKVGEPLWIRCKAIVHNHGFGLTWELEDKALBEGSYFE 300  
Db 241 ECTKLTIDLNOAPQSTLPQLFKVGEPLWIRCKAIVHNHGFGLTWELEDKALBEGSYFE 300  
Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTCSSSKHPSSQALVTILEKGFINATSSOEY 360  
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTCSSSKHPSSQALVTILEKGFINATSSOEY 360  
Qy 361 EIDPYEKFCSVRFKAYPRIRCTWIFSQASFCEQRLGDEGYISKFCDHKNKPGYIFY 420  
Db 361 EIDPYEKFCSVRFKAYPRIRCTWIFSQASFCEQRLGDEGYISKFCDHKNKPGYIFY 420  
Qy 421 AENDDAQFTKMTLNIRKKPQVLANASQASCSDDGYPLPSWTWKCKSDKSPNCTEIP 480  
Db 421 AENDDAQFTKMTLNIRKKPQVLANASQASCSDDGYPLPSWTWKCKSDKSPNCTEIP 480  
Qy 481 EGVNKKANKRVFGQWSSSTLNMSEAGKLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540  
Db 481 EGVNKKANKRVFGQWSSSTLNMSEAGKLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540  
Qy 541 DNISFYATIGLCLPIVVLVILVILCHYKQFRYESQLOMIQVTPGLDNEYFYVDFRDY 600  
Db 541 DNISFYATIGLCLPIVVLVILVILCHYKQFRYESQLOMIQVTPGLDNEYFYVDFRDY 600  
Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOQVAVKMLKEKADSCKEALM 660  
Db 601 DLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOQVAVKMLKEKADSCKEALM 660  
Qy 661 SELKMTHTLGHHDNIVNLGACTLSGVPYLIFFEYCCYGDLLNLYLSKREKPHRTWTEIFK 720  
Db 661 SELKMTHTLGHHDNIVNLGACTLSGVPYLIFFEYCCYGDLLNLYLSKREKPHRTWTEIFK 720  
Qy 721 EHNFSYPTFOAHNSMMPGSRREVQLHPDLQSLGNGNSIHSEDEIYEYENOKRLAE 780  
Db 721 EHNFSYPTFOAHNSMMPGSRREVQLHPDLQSLGNGNSIHSEDEIYEYENOKRLAE 780

QY 781 EDNLVLTFFDILLCFAYQVAKGMEFLKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840  
DB 781 EDNLVLTFFDILLCFAYQVAKGMEFLKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840  
QY 841 SDSSVYVRGNARLPVKWMAPESEFEGYITKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900  
DB 841 SDSSVYVRGNARLPVKWMAPESEFEGYITKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900  
QY 901 FYKLIOGFKMEOPFYATGEGIFVMSQWAFDSRKRPSPNLTSFLGCQLAAEAEACIRT 960  
DB 901 FYKLIOGFKMEOPFYATGEGIFVMSQWAFDSRKRPSPNLTSFLGCQLAAEAEACIRT 960  
QY 961 SYHLPKQAAPOQRGGLRAQSPORQVRIHRERS 992  
DB 961 SYHLPKQAAPOQRGGLRAQSPORQVRIHRERS 992

RESULT 6

US-08-601-891-2  
; Sequence 2, Application US/08601891  
; Patent No. 5747651  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/601,891  
; FILING DATE: 15-FEB-1996  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/977,451  
; FILING DATE: 19-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/679,666  
; FILING DATE: 02-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Felt, Irving N.  
; REGISTRATION NUMBER: 28,601

; REFERENCE/DOCKET NUMBER: LEM-3-7P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405  
; TELEFAX: 212-645-2054  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 992 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-601-891-2  
Query Match 100.0%; Score 5264; DB 1; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MRALAQRSDRRLLLVVLSVMILETVTNQDLPVTKCVLIISHENNGSSAGKPSRYMVRGS 60  
DB 1 MRALAQRSDRRLLLVVLSVMILETVTNQDLPVTKCVLIISHENNGSSAGKPSRYMVRGS 60  
QY 61 PEDLOCTPRQSEGVYEAAATVEAESGSITLQVLATPGDLSCLWVFKHSSLCGCPHF 120  
DB 61 PEDLOCTPRQSEGVYEAAATVEAESGSITLQVLATPGDLSCLWVFKHSSLCGCPHF 120  
QY 121 LQNRGIVSMAILNVTTQAGEYLLHIQSERANYTVLFTVNVVDTQVYVLRPRPYERKMNQ 180  
DB 121 LQNRGIVSMAILNVTTQAGEYLLHIQSERANYTVLFTVNVVDTQVYVLRPRPYERKMNQ 180  
QY 181 DALLCISEGVPEPTVEVWLCSSHRESCKEKGPAVVRKEEVLHFLFGDTRCCARNALGR 240  
DB 181 DALLCISEGVPEPTVEVWLCSSHRESCKEKGPAVVRKEEVLHFLFGDTRCCARNALGR 240  
QY 241 ECTKLFTIDLNOAPOSTLPOLFLKVGPELWIRKCAIHVNHGFGLTWELEDKALEGSGYFE 300  
DB 241 ECTKLFTIDLNOAPOSTLPOLFLKVGPELWIRKCAIHVNHGFGLTWELEDKALEGSGYFE 300  
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSSKHPSQSALVTILEKGFNATSSQBEY 360  
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSSKHPSQSALVTILEKGFNATSSQBEY 360  
QY 361 EIDPYEKFCFSVRFKAYPRICRTWIFSOAQPCEQRLGDCGYSISKFDHKNKPGCEYIF 420  
DB 361 EIDPYEKFCFSVRFKAYPRICRTWIFSOAQPCEQRLGDCGYSISKFDHKNKPGCEYIF 420  
QY 421 AENDDAQFTKMTLNIRKKPQVLANASQASQSDGYPPLPSWTWKKCDKSPNCTEIP 480  
DB 421 AENDDAQFTKMTLNIRKKPQVLANASQASQSDGYPPLPSWTWKKCDKSPNCTEIP 480  
QY 481 EGVWNNKANRKFVGOVSSSTLNMSKAGLLVKCCAYNSMGTSCETIFLNSPGPFPIQ 540  
DB 481 EGVWNNKANRKFVGOVSSSTLNMSKAGLLVKCCAYNSMGTSCETIFLNSPGPFPIQ 540  
QY 541 DNISFYATIGLCLPFIIVLVILVILCHYKFKOPRYESQLOMIQVTPGLDNEYFYVDFRDYEY 600  
DB 541 DNISFYATIGLCLPFIIVLVILVILCHYKFKOPRYESQLOMIQVTPGLDNEYFYVDFRDYEY 600  
QY 601 DLKWEFPRENLFGKVLGSGAFGRVMNATAYGSKTGVSIQVAVKMLKEKADSCKEKALM 660  
DB 601 DLKWEFPRENLFGKVLGSGAFGRVMNATAYGSKTGVSIQVAVKMLKEKADSCKEKALM 660  
QY 661 SELKMMTHLGHNDIVNLGACTLSGPPVYLIFECYCCGDLNLYLRSKREKFRHTWTEIFK 720  
DB 661 SELKMMTHLGHNDIVNLGACTLSGPPVYLIFECYCCGDLNLYLRSKREKFRHTWTEIFK 720  
QY 721 EHNFSYPTFOAHNSNMPGSRVQLHPPLDQLSGFNNGSIHSEDEIEYENOKRLAEDEE 780  
DB 721 EHNFSYPTFOAHNSNMPGSRVQLHPPLDQLSGFNNGSIHSEDEIEYENOKRLAEDEE 780  
QY 781 EDNLVLTFFDILLCFAYQVAKGMEFLKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840  
DB 781 EDNLVLTFFDILLCFAYQVAKGMEFLKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840  
QY 841 SDSSVYVRGNARLPVKWMAPESEFEGYITKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900

Db 841 SDSSVVRGNARLPVKWMAPELFGIYTIKSDVWSYGILLWEIFSLGVNPGIPVDAN 900  
Qy 901 FYKLIQSGFKMEQPYATGEGYFVWQSWAFDSRRKPSFPNLTSLGCOLAAEACIRT 960  
Db 901 FYKLIQSGFKMEQPYATGEGYFVWQSWAFDSRRKPSFPNLTSLGCOLAAEACIRT 960  
Qy 961 SIHLPKQAAPQORGLRAQSPQORQVKIHRERS 992  
Db 961 SIHLPKQAAPQORGLRAQSPQORQVKIHRERS 992

RESULT 7

US-09-021-324-2  
Sequence 2, Application US/09021324  
Patent No. 5912133  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: InClone-Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/021,324  
FILING DATE:  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US/07/977,451  
FILING DATE: 1992-11-19  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:

LENGTH: 992 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-021-324-2  
Query Match 100.0%; Score 5264; DB 2; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MRALAQRSDRRLLLVLSVMILETVTNODLPVVKVLISHENNGSSAGKPSRYMVRGS 60  
Db 1 MRALAQRSDRRLLLVLSVMILETVTNODLPVVKVLISHENNGSSAGKPSRYMVRGS 60  
Qy 61 PEDLOCTPRQSEGTVYEAAATVEAEGSGITLQVQLATPGDLSCLWVFKHSSGCGPHFD 120  
Db 61 PEDLOCTPRQSEGTVYEAAATVEAEGSGITLQVQLATPGDLSCLWVFKHSSGCGPHFD 120  
Qy 121 LQNRGIVSMAILNVTTQAGEYLLHIQSERANYTVLFTVNVTRDTQLYVLRPYFRKMENQ 180  
Db 121 LQNRGIVSMAILNVTTQAGEYLLHIQSERANYTVLFTVNVTRDTQLYVLRPYFRKMENQ 180  
Qy 181 DALLCISEGVPEPTVEWVLCSSHRECKKEEGPAVVRKEEKVLHELFGTDIRCCARNALGR 240  
Db 181 DALLCISEGVPEPTVEWVLCSSHRECKKEEGPAVVRKEEKVLHELFGTDIRCCARNALGR 240  
Qy 241 ECTKLTIDLNAQPOSTLPQLFKYGEPLWIRCKATHVNHGFGLTWELEDKALEEGSYFE 300  
Db 241 ECTKLTIDLNAQPOSTLPQLFKYGEPLWIRCKATHVNHGFGLTWELEDKALEEGSYFE 300  
Qy 301 MSTYSTNTRMIRILLAFVSSVGRNDTGYTSCSSKHPQSALVTILEKGFINATSSQEEY 360  
Db 301 MSTYSTNTRMIRILLAFVSSVGRNDTGYTSCSSKHPQSALVTILEKGFINATSSQEEY 360  
Qy 361 EIDPYEKFCFSVRKAYPRICTWIFSQASFCEQRLGLEDGYSISKFDHKNKPGYIFY 420  
Db 361 EIDPYEKFCFSVRKAYPRICTWIFSQASFCEQRLGLEDGYSISKFDHKNKPGYIFY 420  
Qy 421 AENDDAQFTKMTLNIRKPKQVLANASQASCSDDGYPLPSWTWKKSDKSPNCTEIP 480  
Db 421 AENDDAQFTKMTLNIRKPKQVLANASQASCSDDGYPLPSWTWKKSDKSPNCTEIP 480  
Qy 481 EGVWNNKANRKYFGOWVSSSTLNMSSEAGLLVKCCAYNSMGTSCTETFLNSGPPFFIQ 540  
Db 481 EGVWNNKANRKYFGOWVSSSTLNMSSEAGLLVKCCAYNSMGTSCTETFLNSGPPFFIQ 540  
Qy 541 DNISFYATIGLCLPFIVLVILVILCHIKYKQFRYESOLQMTQVTPGLDNEYFYVDFRDYEY 600  
Db 541 DNISFYATIGLCLPFIVLVILVILCHIKYKQFRYESOLQMTQVTPGLDNEYFYVDFRDYEY 600  
Qy 601 DLKWEFPRENLEFPGKVLGSGAGFRVNNATAYGISTGVSIOQVAVKMLKEKADSCKEALM 660  
Db 601 DLKWEFPRENLEFPGKVLGSGAGFRVNNATAYGISTGVSIOQVAVKMLKEKADSCKEALM 660  
Qy 661 SELKMMTHLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLAYLRSKRKFHRTWTETIFK 720  
Db 661 SELKMMTHLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLAYLRSKRKFHRTWTETIFK 720  
Qy 721 EHNFSYPTFOAHNSNMPGSRREVQLHPDLQSLGFGNSIHSDEIEYENQKRLAEEEE 780  
Db 721 EHNFSYPTFOAHNSNMPGSRREVQLHPDLQSLGFGNSIHSDEIEYENQKRLAEEEE 780  
Qy 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHVKKVVICDFGLARDIL 840  
Db 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHVKKVVICDFGLARDIL 840  
Qy 841 SDSSVVRGNARLPVKWMAPELFGIYTIKSDVWSYGILLWEIFSLGVNPGIPVDAN 900  
Db 841 SDSSVVRGNARLPVKWMAPELFGIYTIKSDVWSYGILLWEIFSLGVNPGIPVDAN 900  
Qy 901 FYKLIQSGFKMEQPYATGEGYFVWQSWAFDSRRKPSFPNLTSLGCOLAAEACIRT 960  
Db 901 FYKLIQSGFKMEQPYATGEGYFVWQSWAFDSRRKPSFPNLTSLGCOLAAEACIRT 960



TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 992 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-05401-2

Query Match 100.0%; Score 5264; DB 5; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALAQRSDRRLLLVLSVMILETVTNQDLPVTKVLIHENNGSSAGKPSRYMVRGS 60  
DB 1 MRALAQRSDRRLLLVLSVMILETVTNQDLPVTKVLIHENNGSSAGKPSRYMVRGS 60  
QY 61 PEDLOCTPRROSEGVTEAATVEAESGSLTQVOLTATPGDLSCLWPKHSSLCQPHFD 120  
DB 61 PEDLOCTPRROSEGVTEAATVEAESGSLTQVOLTATPGDLSCLWPKHSSLCQPHFD 120  
QY 121 LQNRGIVSMALINVTETQAGEYLLHIQSERANYTVLFTVNVNVDLTQVLRPRYFRKMEQ 180  
DB 121 LQNRGIVSMALINVTETQAGEYLLHIQSERANYTVLFTVNVNVDLTQVLRPRYFRKMEQ 180  
QY 181 DALLCISEGVPEPTVWVLCSSHRESCKEKGPAVVRKEEYVHLFGTDIRCCARNALGR 240  
DB 181 DALLCISEGVPEPTVWVLCSSHRESCKEKGPAVVRKEEYVHLFGTDIRCCARNALGR 240  
QY 241 ECTKLFTIDLNOAQPOSTLPOLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300  
DB 241 ECTKLFTIDLNOAQPOSTLPOLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300  
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPQSALVTILEKGFNATSSOEY 360  
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPQSALVTILEKGFNATSSOEY 360  
QY 361 EIDPYEKFCFSVRKAYPRIRCTWIFSOAGFPCEQRLGEGYSISKCDHKNKPGEXIFY 420  
DB 361 EIDPYEKFCFSVRKAYPRIRCTWIFSOAGFPCEQRLGEGYSISKCDHKNKPGEXIFY 420  
QY 421 AENDDAQFTKMTNIRKKQVLANASASQSCSDGYPPLPSWTWKCKSDKSPNCTEIP 480  
DB 421 AENDDAQFTKMTNIRKKQVLANASASQSCSDGYPPLPSWTWKCKSDKSPNCTEIP 480  
QY 481 EGVNKKANKRVFGQWYSSSTLNNSEAGKLLVKCCAYNSMGTSCTETFLNSPGPFPIQ 540  
DB 481 EGVNKKANKRVFGQWYSSSTLNNSEAGKLLVKCCAYNSMGTSCTETFLNSPGPFPIQ 540  
QY 541 DNISFYATIGLCLPFIIVLVILVICHKKQFRIESQLOMQIOVTGPDNEFYVDFRDY 600  
DB 541 DNISFYATIGLCLPFIIVLVILVICHKKQFRIESQLOMQIOVTGPDNEFYVDFRDY 600  
QY 601 DLKWEFPRENLEFGKVLGSAFGVMNATAYGSKTGVSIOVAVKMLKEKADSCKEKALM 660  
DB 601 DLKWEFPRENLEFGKVLGSAFGVMNATAYGSKTGVSIOVAVKMLKEKADSCKEKALM 660  
QY 661 SELKMMTHLGHHDNVLNGLACTLSGPVYLFEYCCYGDLLNLRKREKPHRTWTEIFK 720  
DB 661 SELKMMTHLGHHDNVLNGLACTLSGPVYLFEYCCYGDLLNLRKREKPHRTWTEIFK 720  
QY 721 EHNFSYPTFOAHNSNMPGSRVQLHPPDLQSLGFGNNSHSEDEYENOKRLAEDEE 780  
DB 721 EHNFSYPTFOAHNSNMPGSRVQLHPPDLQSLGFGNNSHSEDEYENOKRLAEDEE 780  
QY 781 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVYKICDFGLARDIL 840  
DB 781 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVYKICDFGLARDIL 840  
QY 841 SDSSYVVRGNARLPVKWMAPESLPEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900  
DB 841 SDSSYVVRGNARLPVKWMAPESLPEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900

DB 841 SDSSYVVRGNARLPVKWMAPESLPEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900  
QY 901 FYKLIQSGFKMEQPFYATGEGYFVWQSWAFDSRKRPSFPNLTSLGCOLABAEACIRT 960  
DB 901 FYKLIQSGFKMEQPFYATGEGYFVWQSWAFDSRKRPSFPNLTSLGCOLABAEACIRT 960  
QY 961 SIHLPKQAAPQOGRGLRAQSPQORVKIHRERS 992  
DB 961 SIHLPKQAAPQOGRGLRAQSPQORVKIHRERS 992

RESULT 10

PCT-US92-09893-2  
Sequence 2, Application PC/TUS9209893  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Imclone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/09893  
FILING DATE: 19921116  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N. 28, 601  
REGISTRATION NUMBER: LEM-3-7PT  
REFERENCE/DOCKET NUMBER: LEM-3-7PT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 992 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-09893-2

Query Match 100.0%; Score 5264; DB 5; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALAQRSDRRLLLVLSVMILETVTNQDLPVTKVLIHENNGSSAGKPSRYMVRGS 60  
DB 1 MRALAQRSDRRLLLVLSVMILETVTNQDLPVTKVLIHENNGSSAGKPSRYMVRGS 60  
QY 61 PEDLOCTPRROSEGVTEAATVEAESGSLTQVOLTATPGDLSCLWPKHSSLCQPHFD 120  
DB 61 PEDLOCTPRROSEGVTEAATVEAESGSLTQVOLTATPGDLSCLWPKHSSLCQPHFD 120  
QY 121 LQNRGIVSMALINVTETQAGEYLLHIQSERANYTVLFTVNVNVDLTQVLRPRYFRKMEQ 180  
DB 121 LQNRGIVSMALINVTETQAGEYLLHIQSERANYTVLFTVNVNVDLTQVLRPRYFRKMEQ 180  
QY 181 DALLCISEGVPEPTVWVLCSSHRESCKEKGPAVVRKEEYVHLFGTDIRCCARNALGR 240  
DB 181 DALLCISEGVPEPTVWVLCSSHRESCKEKGPAVVRKEEYVHLFGTDIRCCARNALGR 240  
QY 241 ECTKLFTIDLNOAQPOSTLPOLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300  
DB 241 ECTKLFTIDLNOAQPOSTLPOLFLKVGEPWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300

QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVTTILEKGFINATSSOEY 360  
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVTTILEKGFINATSSOEY 360  
QY 361 EIDPYEKFCFSVRKAYPRIRCTWIFSOASFPCEQGLDGYISKFCDHKNKPGYIFY 420  
DB 361 EIDPYEKFCFSVRKAYPRIRCTWIFSOASFPCEQGLDGYISKFCDHKNKPGYIFY 420  
QY 421 AENDDAQFTKMTLNIRKPKQVLANASASQSCSDGYPLPSWTWKCKSDKSPNCTEIP 480  
DB 421 AENDDAQFTKMTLNIRKPKQVLANASASQSCSDGYPLPSWTWKCKSDKSPNCTEIP 480  
QY 481 EGVNKKANKRVFGQWVSSTLNMSAGKGLLVKCCAYNSMGTSCTETIFLNSGPPFFIQ 540  
DB 481 EGVNKKANKRVFGQWVSSTLNMSAGKGLLVKCCAYNSMGTSCTETIFLNSGPPFFIQ 540  
QY 541 DNISFYATIGLCLPFIIVLVILVILCHYKQFRIESOLOMIQVTPLDNEYFYVDFRDY 600  
DB 541 DNISFYATIGLCLPFIIVLVILVILCHYKQFRIESOLOMIQVTPLDNEYFYVDFRDY 600  
QY 601 DLKWEPPRENLEFGKVLGSGAFGRVNMATAYGISKTSQVSIQVAVKMLKEKADSCKEALM 660  
DB 601 DLKWEPPRENLEFGKVLGSGAFGRVNMATAYGISKTSQVSIQVAVKMLKEKADSCKEALM 660  
QY 661 SELKMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKFRHTWTIEFK 720  
DB 661 SELKMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKFRHTWTIEFK 720  
QY 721 EHNFSYPTFOAHSSNMPGSRREVLPDQLDLSGFGNSIHSEDEIEYENKRLAEEB 780  
DB 721 EHNFSYPTFOAHSSNMPGSRREVLPDQLDLSGFGNSIHSEDEIEYENKRLAEEB 780  
QY 781 EDNLVLTEDLLCFAYAKGMEFLFSCVHRDLAARNVLVTHGKVKYKICDFGLARDIL 840  
DB 781 EDNLVLTEDLLCFAYAKGMEFLFSCVHRDLAARNVLVTHGKVKYKICDFGLARDIL 840  
QY 841 SDSSVVRGNARLPVKWAPESLFGIYTKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900  
DB 841 SDSSVVRGNARLPVKWAPESLFGIYTKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900  
QY 901 FYKLQSGFKMQPYATGEGYIFVMSQWAFDSRKRPSFNLTSLGCOLAABEACIRT 960  
DB 901 FYKLQSGFKMQPYATGEGYIFVMSQWAFDSRKRPSFNLTSLGCOLAABEACIRT 960  
QY 961 SIHLPKQAAPQORGLRAQSPORQVKIHRERS 992  
DB 961 SIHLPKQAAPQORGLRAQSPORQVKIHRERS 992

## RESULT 11

US-08-222-299-2  
; Sequence 2, Application US/08222299  
; Patent No. 5635388  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, Brian D.  
; APPLICANT: Broz, Susan D.  
; APPLICANT: Matthews, William  
; APPLICANT: Zeigler, Francis C.  
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THEREOF  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)

## CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/222,299  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hasak, Janet E.  
; REGISTRATION NUMBER: 28,616  
; REFERENCE/DOCKET NUMBER: 879  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1896  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1000 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-222-299-2

Query Match 96.9%; Score 5102; DB 1; Length 1000;

Best Local Similarity 96.7%; Pred. No. 0;

Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;

QY 1 MRALAQRSDRRLLLVLSVMIETVNTQDLPVICKVLISHENNGSSAGKPSRYRMVRS 60  
DB 1 MRALAQRSDRRLLLVLSVMIETVNTQDLPVICKVLISHENNGSSAGKPSRYRMVRS 60  
QY 61 PEDLOCTPRRQSEGTVYEATVEAEGSSITLQVQLATPGDLSCLVWFKHSSLGCCPHFD 120  
DB 61 PEDLOCTPRRQSEGTVYEATVEAEGSSITLQVQLATPGDLSCLVWFKHSSLGCCPHFD 120  
QY 121 LQNRGIVSMALNVETQAGEYLLHIQSERANTVLTFTVNVRTQYLVLRPFRKMEQ 180  
DB 121 LQNRGIVSMALNVETQAGEYLLHIQSERANTVLTFTVNVRTQYLVLRPFRKMEQ 180  
QY 181 DALLCISEGVPEPTVEWILCSSHRESCKEGPAVVRKEEVLHELFGTDIRCCARNALGR 240  
DB 181 DALLCISEGVPEPTVEWILCSSHRESCKEGPAVVRKEEVLHELFGTDIRCCARNALGR 240  
QY 241 ECTKFTIDLNAQPSQTLPLKVGEPWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300  
DB 241 ECTKFTIDLNAQPSQTLPLKVGEPWIRCKAIHVNHGFGLTWELEDKALEEGSYFE 300  
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVTTILEKGFINATSSOEY 360  
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVTTILEKGFINATSSOEY 360  
QY 361 EIDPYEKFCFSVRKAYPRIRCTWIFSOASFPCEQGLDGYISKFCDHKNKPGYIFY 420  
DB 361 EIDPYEKFCFSVRKAYPRIRCTWIFSOASFPCEQGLDGYISKFCDHKNKPGYIFY 420  
QY 421 AENDDAQFTKMTLNIRKPKQVLANASASQSCSDGYPLPSWTWKCKSDKSPNCTEIP 480  
DB 421 AENDDAQFTKMTLNIRKPKQVLANASASQSCSDGYPLPSWTWKCKSDKSPNCTEIP 480  
QY 481 EGVNKKANKRVFGQWVSSTLNMSAGKGLLVKCCAYNSMGTSCTETIFLNSGPPFFIQ 540  
DB 481 EGVNKKANKRVFGQWVSSTLNMSAGKGLLVKCCAYNSMGTSCTETIFLNSGPPFFIQ 540  
QY 541 DNISFYATIGLCLPFIIVLVILVILCHYKQFRIESOLOMIQVTPLDNEYFYVDFRDY 600  
DB 541 DNISFYATIGLCLPFIIVLVILVILCHYKQFRIESOLOMIQVTPLDNEYFYVDFRDY 600  
QY 601 DLKWEPPRENLEFGKVLGSGAFGRVNMATAYGISKTSQVSIQVAVKMLKEKADSCKEALM 660  
DB 601 DLKWEPPRENLEFGKVLGSGAFGRVNMATAYGISKTSQVSIQVAVKMLKEKADSCKEALM 660  
QY 661 SELKMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKFRHTWTIEFK 720  
DB 661 SELKMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKFRHTWTIEFK 720

QY 721 EHNFSYPTFOAHNSMPSGSRVQLHPPDLQSLGFGNGSIHSEDEIEYENOKRLAEED 780  
Db 721 EHNFSYPTFOAHNSMPSGSRVQLHPPDLQSLGFGNGSIHSEDEIEYENOKRLAEED 780  
QY 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGKVKYKICDFGLARDIL 840  
Db 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGKVKYKICDFGLARDIL 840  
QY 841 SDSSVYVRGNARLPVKWMAPELFEIYIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900  
Db 841 SDSSVYVRGNARLPVKWMAPELFEIYIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900  
QY 901 FYKLIQSGFKMEQPFYATGFIYFVMSQWAFDSRRKPSFPNLTSLGCOLAEAEACIRT 960  
Db 901 FYKLIQSGFKMEQPFYATGFIYFVMSQWAFDSRRKPSFPNLTSLGCOLAEAEAMYN 960  
QY 961 -----SIHLPKQAAPQGRGLRAQSPQKVIHRS 992  
Db 961 MGGNVPEHPSIYQNRRLSREAGS-EPPSPQAQVKIHR 1000

RESULT 12  
US-08-434-878-2  
; Sequence 2, Application US/08434878  
; Patent No. 5997865  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, Brian D.  
; APPLICANT: Broz, Susan D.  
; APPLICANT: Matthews, William  
; APPLICANT: Zeigler, Francis C.  
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THEREOF  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/434,878  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hasak, Janet E.  
; REGISTRATION NUMBER: 28,616  
; REFERENCE/DOCKET NUMBER: 879  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1896  
; TELEFAX: 415/952-9881  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1000 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-434-878-2

Query Match 96.98; Score 5102; DB 2; Length 1000;  
Best Local Similarity 96.78; Pred. No. 0;  
Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;  
QY 1 MRALAQRSDRLLLVLSVMILETVTNODLPVICKVLISHENNGSSACKPSSYRMVRGS 60  
Db 1 MRALAQRSDRLLLVLSVMILETVTNODLPVICKVLISHENNGSSACKPSSYRMVRGS 60  
QY 61 PEDLOCTPRQSEGTVYEAAATVEAESGSIITLQVQLATPGDLSCLVWFKHSSJGCOHPFD 120  
Db 61 PEDLOCTPRQSEGTVYEAAATVEAESGSIITLQVQLATPGDLSCLVWFKHSSJGCOHPFD 120  
QY 121 LQNRGIVSMAILNVETQAGEYLLHIQSERANYTVLFTVNVDRDTQLYVLRPRYFRKMEQ 180  
Db 121 LQNRGIVSMAILNVETQAGEYLLHIQSERANYTVLFTVNVDRDTQLYVLRPRYFRKMEQ 180  
QY 181 DALLCISEGVPPTVEWVLCSSHRSCKEKGPAVVRKEKVLHELFGTDIRCCARNALGR 240  
Db 181 DALLCISEGVPPTVEWVLCSSHRSCKEKGPAVVRKEKVLHELFGTDIRCCARNALGR 240  
QY 241 ECTKLTIDLNAPOSTLPOLFKVGEPLWIRCKATHVNHGFGLTWELEDKALEESYFE 300  
Db 241 ECTKLTIDLNAPOSTLPOLFKVGEPLWIRCKATHVNHGFGLTWELEDKALEESYFE 300  
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTSCSSKHPQSALVTILEKGFINATSSQEEY 360  
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTSCSSKHPQSALVTILEKGFINATSSQEEY 360  
QY 361 EIDPYEKFCFSVRFRKAYPRIRCTWIFSQASFCEQRLGLEDGYSISKFOHKNKPGYIFY 420  
Db 361 EIDPYEKFCFSVRFRKAYPRIRCTWIFSQASFCEQRLGLEDGYSISKFOHKNKPGYIFY 420  
QY 421 AENDDAQFTKMETLNIRKKPOVLANASASQASCSGYPPLPSWTWKKCSDKSPNCTEIP 480  
Db 421 AENDDAQFTKMETLNIRKKPOVLANASASQASCSGYPPLPSWTWKKCSDKSPNCTEIP 480  
QY 481 EGVWKNKANKRVFGOWSSSTLNMSSEAGKGLLVKCCAYNSMGTSCETIFLNSGPPFFIQ 540  
Db 481 EGVWKNKANKRVFGOWSSSTLNMSSEAGKGLLVKCCAYNSMGTSCETIFLNSGPPFFIQ 540  
QY 541 DNISFYATIGLCLPFIIVLVILICHKYKQRYESOLQMIQVTPGLDNEYFYVDFRDY 600  
Db 541 DNISFYATIGLCLPFIIVLVILICHKYKQRYESOLQMIQVTPGLDNEYFYVDFRDY 600  
QY 601 DLKWEFPRENLEFGKVLGSGAGRVNNAAYGISTKGVSIQVAVKMLKADSCKEALM 660  
Db 601 DLKWEFPRENLEFGKVLGSGAGRVNNAAYGISTKGVSIQVAVKMLKADSCKEALM 660  
QY 661 SELKMMTHLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLNLYRSKRKFHRTWTIEFK 720  
Db 661 SELKMMTHLGHHDNIVNLGACTLSGPVYLIFEYCCYGDLLNLYRSKRKFHRTWTIEFK 720  
QY 721 EHNFSYPTFOAHNSMPSGSRVQLHPPDLQSLGFGNGSIHSEDEIEYENOKRLAEED 780  
Db 721 EHNFSYPTFOAHNSMPSGSRVQLHPPDLQSLGFGNGSIHSEDEIEYENOKRLAEED 780  
QY 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGKVKYKICDFGLARDIL 840  
Db 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGKVKYKICDFGLARDIL 840  
QY 841 SDSSVYVRGNARLPVKWMAPELFEIYIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900  
Db 841 SDSSVYVRGNARLPVKWMAPELFEIYIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900  
QY 901 FYKLIQSGFKMEQPFYATGFIYFVMSQWAFDSRRKPSFPNLTSLGCOLAEAEACIRT 960  
Db 901 FYKLIQSGFKMEQPFYATGFIYFVMSQWAFDSRRKPSFPNLTSLGCOLAEAEAMYN 960  
QY 961 -----SIHLPKQAAPQGRGLRAQSPQKVIHRS 992  
Db 961 MGGNVPEHPSIYQNRRLSREAGS-EPPSPQAQVKIHR 1000

RESULT 13  
PCT-US95-03718-2  
; Sequence 2, Application PC/TUS9503718  
; GENERAL INFORMATION:  
; APPLICANT: GENENTECH, INC.  
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES THEREOF



NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/03718  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Wendy M. Lee  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 879PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1000 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
PCT-US95-03718-2

Query Match 96.9%; Score 5102; DB 5; Length 1000;  
Best Local Similarity 96.7%; Pred. No. 0;  
Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;  
Qy 1 MRALAQRSDRLLLLVLSVMILETVTNODLPVKVLISHENNGSSAGKPSYRMVRS 60  
Db 1 MRALAQRSDRLLLLVLSVMILETVTNODLPVKVLISHENNGSSAGKPSYRMVRS 60  
Qy 61 PEDLOCTPRROSEGTVYEAATVEAESGSILOVOLATPGDLSCLWVFKHSSIGCOPHD 120  
Db 61 PEDLOCTPRROSEGTVYEAATVEAESGSILOVOLATPGDLSCLWVFKHSSIGCOPHD 120  
Qy 121 LQNRGIVSMALNVTTQAGEYLLHQSRANVTVLFTVNVRTQLYVLRPPYFRKMENQ 180  
Db 121 LQNRGIVSMALNVTTQAGEYLLHQSRANVTVLFTVNVRTQLYVLRPPYFRKMENQ 180  
Qy 181 DALLCISEGVPEPTVYEWVLCSSHRESCKEEGPAVVRKEEVLHFGTDIRCCARNALGR 240  
Db 181 DALLCISEGVPEPTVYEWVLCSSHRESCKEEGPAVVRKEEVLHFGTDIRCCARNALGR 240  
Qy 241 ECTKLTIDLNAQPOSTLPOLKLVGEPLWIRCKAIHVNHGGLTWELDKALEEGSYFE 300  
Db 241 ECTKLTIDLNAQPOSTLPOLKLVGEPLWIRCKAIHVNHGGLTWELDKALEEGSYFE 300  
Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTSCSSKHPSOSALVTILEKGFINATSSQEEY 360  
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTSCSSKHPSOSALVTILEKGFINATSSQEEY 360  
Qy 361 EIDPYEKFCSVRFKAYPIRITWIFSQAFCEQRLGEGYSISKFDHKNKPGYIFY 420  
Db 361 EIDPYEKFCSVRFKAYPIRITWIFSQAFCEQRLGEGYSISKFDHKNKPGYIFY 420  
Qy 421 AENDDAQFTKMTLNIRKKPQVLANASQASCSGYPPLPSTWTKKSDKSPNCTEIP 480  
Db 421 AENDDAQFTKMTLNIRKKPQVLANASQASCSGYPPLPSTWTKKSDKSPNCTEIP 480  
Qy 481 EGVWNNKANRKYFGQWVSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540

Db 481 EGVWNNKANRKYFGQWVSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPGPPFIQ 540  
Qy 541 DNISFYATIGLCLPFIIVLVILVILCHIKYKOPRYESOLOMIQVTGPDNDNEYFYVDFRDY 600  
Db 541 DNISFYATIGLCLPFIIVLVILVILCHIKYKOPRYESOLOMIQVTGPDNDNEYFYVDFRDY 600  
Qy 601 DLKWEFFPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOQAVKMLKEKADSCKEALM 660  
Db 601 DLKWEFFPRENLEFGKVLGSGAFGRVMNATAYGISKTGVSIOQAVKMLKEKADSCKEALM 660  
Qy 661 SELKMMTHLGHHDNIIVNLGACTLSGPVYLIFEYCCYCGDLLNLYLSKREKFRHTWTEIFK 720  
Db 661 SELKMMTHLGHHDNIIVNLGACTLSGPVYLIFEYCCYCGDLLNLYLSKREKFRHTWTEIFK 720  
Qy 721 EHNFSYPTFOAHSSNMPGSRVOLHPDLQDLSGNGNSIHSEDEIEYENOKRLAESEE 780  
Db 721 EHNFSYPTFOAHSSNMPGSRVOLHPDLQDLSGNGNSIHSEDEIEYENOKRLAESEE 780  
Qy 781 EDNLVLTEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLTGHGVKVKICDFGLARDIL 840  
Db 781 EDNLVLTEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLTGHGVKVKICDFGLARDIL 840  
Qy 841 SDSYVVRGNARLPVKWMAPELSLFEIGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 900  
Db 841 SDSYVVRGNARLPVKWMAPELSLFEIGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 900  
Qy 901 FYKLIQSGFKMEQPPFYATEGIYFVMSQWAFDSRKRPSFNLTSLGCGQLAEAEACIRT 960  
Db 901 FYKLIQSGFKMEQPPFYATEGIYFVMSQWAFDSRKRPSFNLTSLGCGQLAEAEAMYNQ 960  
Qy 961 -----SIHLPKQAAPQORGLRAQSPQROVKIHRERS 992  
Db 961 MGNVPEHPSIYQNRRLPSREAGS-EPPSPQAOVKIHRERS 1000  
RESULT 14  
US-08-222-299-4  
Sequence 4, Application US/08222299  
Patent No. 5635388  
GENERAL INFORMATION:  
APPLICANT: Bennett, Brian D.  
APPLICANT: Broz, Susan D.  
APPLICANT: Matthews, William  
APPLICANT: Zeigler, Francis C.  
TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/222,299  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 879  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-222-299-4

Query Match 84.3%; Score 4436.5; DB 1; Length 993;  
Best Local Similarity 84.0%; Pred. No. 0;  
Matches 836; Conservative 59; Mismatches 87; Indels 13; Gaps 4;

Qy 1 MRALQSRDRLLLVLSVMIETVNTODLPVVKVLSHNNSSAGKPSRYRWGRS 60  
Db 1 MPALA-RDGGQLPLLWVFSAMIFGTITNODLPVVKVLSHNNSSAGKPSRYRWGRS 59

Qy 61 PEDLOCTPRROSECTVYEAATVEVAESGSITLOVOLATPGDLSCLVWFKHSLGCGPHD 120  
Db 60 PEDLGALRPOSSGTVYEAATVEVAESGSITLOVLDPAGNISCLVWFKHSLGCGPHD 119

Qy 121 LQNRGIVSMALNTVTOAGEYLLHIQSERANYTLVTVNVRDITOLYVLRPRYFRKMENQ 180  
Db 120 LQNRGVSMVLKMTETQAGEYLLFIQSEATNYTLFVSTINLTLLYLRPRYFRKMENQ 179

Qy 181 DALLCISEGVPEPTVWVLCSSHRESCKEKPVAVYRKEEVLHFGTDIRCCARNALGR 240  
Db 180 DALVCISESVPEPTVWVLCDSQGESCKEESPAVVKKEEVLHFGMDIRCCARNELGR 239

Qy 301 MSTYSTNRTMIRILAFVSSVGRNDTYTCTSSKHPQSOSALVTILEKGFNATSSQEEY 360  
Db 300 MSTYSTNRTMIRILAFVSSVARNDTYTCTSSKHPQSOSALVTIVEKGFNATNSSDY 359

Qy 361 EIDPYEFCFSVREKAYPRICRTWIFSOAPCEQGLDGYISKFCDHKNKPGCEYIF 420  
Db 360 EIDQYEEFCFSVREKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNHHKQPGCEYIF 419

Qy 421 AENDDAQFTKMTNIRKKPOVLNANASQASCSGDPVLPSTWTKKSDKSPNCTEIP 480  
Db 420 AENDDAQFTKMTNIRKKPOVLNANASQASCSGDPVLPSTWTKKSDKSPNCTEIT 479

Qy 481 EGVNKKANRVFCQWSSSTPLNNSEAGLLVKKCAVNSNGTSCETIFLNSPGPFPIQ 540  
Db 480 EGVNKKANRVFCQWSSSTPLNNSEALVKKCAVNSNGTSCETIFLNSPGPFPIQ 539

Qy 541 DNISFYATIGLCLPFIYVLIIVLICHKKQFRYESQLOMIQVTPDNEFYVDFRDY 600  
Db 540 DNISFYATIGVCLLFIYVLTLLICHKKQFRYESQLOMQVGTSSDNEYFYVDFREY 599

Qy 601 DLKWEFPRENLEFGKVLGSGAFGRVNNATAYGSKTGVSIOVAVKMLKEKADSCKEALM 660  
Db 600 DLKWEFPRENLEFGKVLGSGAFGRVNNATAYGSKTGVSIOVAVKMLKEKADSSEREALM 659

Qy 661 SELKMTHLGHHDHNVLLGACTLSGPVYLIFCYCCYGDLLNLYRSKREKFRHTWTEIF 720  
Db 660 SELKMTQLGSHENIVNLLGACTLSGPVYLIFCYCCYGDLLNLYRSKREKFRHTWTEIF 719

Qy 721 EHNFSYPTFOAHNSNMPGSRVQLHPPLDQLSGFNGNSIHSEDEIEYENOKRLAE 780  
Db 720 EHNFSYPTFOAHNSNMPGSRVQLHPPLDQLSGFNGNSIHSEDEIEYENOKRL--EE 777

Qy 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840  
Db 778 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIM 837

Qy 841 SDSYVVRGNARLPVKWAPPSLEGIYTIKSDVWSVIGILLWEIFSLGVNYPGIPVDAN 900  
Db 838 SDSYVVRGNARLPVKWAPPSLEGIYTIKSDVWSVIGILLWEIFSLGVNYPGIPVDAN 897

Qy 901 FYKLIQSGFKWEQPFYATEGIFYVMQSCWAFDSRKRPSFNLTSFLGCOLADAEAEACIRT 960  
Db 900 FYKLIQSGFKWEQPFYATEGIFYVMQSCWAFDSRKRPSFNLTSFLGCOLADAEAEACIRT 956

Db 898 FYKLIQSGFKWEQPFYATEGIFYVMQSCWAFDSRKRPSFNLTSFLGCOLADAEAEANYQ- 956

Qy 961 SIHLPKQAP-----QORGGLRAQSPQOVK 986

Db 957 NVDPVSECPHTYQNRPRPSREMDLGLLSPQAQVE 991

RESULT 15  
US-08-434-878-4  
Sequence 4, Application US/08434878  
Patent No. 5997865  
GENERAL INFORMATION:  
APPLICANT: Bennett, Brian D.  
APPLICANT: Broz, Susan D.  
APPLICANT: Matthews, William  
APPLICANT: Zeigler, Francis C.  
TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/434,878  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 879  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-434-878-4

Query Match 84.3%; Score 4436.5; DB 2; Length 993;  
Best Local Similarity 84.0%; Pred. No. 0;  
Matches 836; Conservative 59; Mismatches 87; Indels 13; Gaps 4;

Qy 1 MRALQSRDRLLLVLSVMIETVNTODLPVVKVLSHNNSSAGKPSRYRWGRS 60  
Db 1 MPALA-RDGGQLPLLWVFSAMIFGTITNODLPVVKVLSHNNSSAGKPSRYRWGRS 59

Qy 61 PEDLOCTPRROSECTVYEAATVEVAESGSITLOVOLATPGDLSCLVWFKHSLGCGPHD 120  
Db 60 PEDLGALRPOSSGTVYEAATVEVAESGSITLOVLDPAGNISCLVWFKHSLGCGPHD 119

Qy 121 LQNRGIVSMALNTVTOAGEYLLHIQSERANYTLVTVNVRDITOLYVLRPRYFRKMENQ 180  
Db 120 LQNRGVSMVLKMTETQAGEYLLFIQSEATNYTLFVSTINLTLLYLRPRYFRKMENQ 179

Qy 181 DALLCISEGVPEPTVWVLCSSHRESCKEKPVAVYRKEEVLHFGTDIRCCARNALGR 240  
Db 180 DALVCISESVPEPTVWVLCDSQGESCKEESPAVVKKEEVLHFGMDIRCCARNELGR 239

Qy 241 ECTKLFTTDLNQAPOSTLPQLFLKVGPELWIRCKAIHVNHGFLTWELDEKALEEGSYFE 300

Db 240 ECTRLFTIDLNQTTQLPOLFLKAVGEPWIRCKAVHVNHGFGLTWELENKALEEGNYFE 299  
Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360  
Db 300 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSQSALVTIVEKGFINATNSSEY 359  
Qy 361 EIDPYEKFCFSVREKAYPRICRTWIFSOASPPCQORGLDGYISKFCDHKNKPGYEYIF 420  
Db 360 EIDQYEEFCFSVREKAYPOICRTWTFSRKSPCQKGLDNGYSISKFCNHHKHQGEYIFH 419  
Qy 421 AENDDAOFTKMTNIRKPOVLANASASQSCSDGYPLPSWTWKKCSDKSPNCTEIP 480  
Db 420 AENDDAOFTKMTNIRKPOVLAEASASQSCSDGYPLPSWTWKKCSDKSPNCTEIT 479  
Qy 481 EGVNKNRNVFGQWSSSFLNMSSEAGKLLVCCAYNSMGTSCETIFLNSPGPFPIQ 540  
Db 480 EGVNKNRNVFGQWSSSFLNMSSEAIKGLVKCCAYNSLGTSCETILLNSPGPFPIQ 539  
Qy 541 DNISFYATIGLCLPIVVLIVLICHYKKQFRIESOLOMIQVTGPDNEYFYVDFRDY 600  
Db 540 DNISFYATIGVCLLPIVVLTLICHYKKQFRIESOLOMVQVTGSSDNEYFYVDFREY 599  
Qy 601 DLKWEFFPRENLEFGKVLGSGAFGRYMNATAYGISKTGYSIQVAVKMLKEKADSCKEALM 660  
Db 600 DLKWEFFPRENLEFGKVLGSGAFGRYMNATAYGISKTGYSIQVAVKMLKEKADSSREALM 659  
Qy 661 SELKMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYRSKREKPHRTWTEIFK 720  
Db 660 SELKMTQLGSHENIVNLLGACTLSGPVYLIFEYCCYGDLLNLYRSKREKPHRTWTEIFK 719  
Qy 721 EHNFSYPTFOAHNSNMPGSGREVOLHPPLDOLSGFNHSEDEIEYENOKRLAEDEE 780  
Db 720 EHNFSYPTFOAHNSNMPGSGREVOLHPDQISGLHNSFHSEDEIEYENOKRL-EE 777  
Qy 781 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840  
Db 778 EDNLVLTEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHGVKVICDFGLARDIM 837  
Qy 841 SDSNVVRGNARLPVKWMAPESLFEGYITIKSDVNSYGILLWEIFSLGVNYPGIPVDAN 900  
Db 838 SDSNVVRGNARLPVKWMAPESLFEGYITIKSDVNSYGILLWEIFSLGVNYPGIPVDAN 897  
Qy 901 FYKLQSGFKMEQPFYATEGIYVWOSWAFDSRKRPSPNLTSLGCOLAEAEACIRT 960  
Db 898 FYKLQSGFKMEQPFYATEGIYVWOSWAFDSRKRPSPNLTSLGCOLAEAEAMQ- 956  
Qy 961 STHLPKQAAAP-----QORGGRLRAQSPQOVK 986  
Db 957 NVDGPVSECPHTYQNRPRPSRENDLGLLSPOAQVE 991

Search completed: May 27, 2003, 14:36:15  
Job time : 28.9879 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:34:40 ; Search time 24.4877 seconds  
(without alignments)  
4017.505 Million cell updates/sec

Title: US-09-919-408-2

Perfect score: 5264

Sequence: 1 MRALAQRDRRLLLVVLVS.....RGGLRAQSPQVQKIHRS 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5264	100.0	992	10	US-09-919-408-2
2	5264	100.0	992	10	US-09-872-136-2
3	4429.5	84.1	993	10	US-09-919-408-4
4	4429.5	84.1	993	10	US-09-872-136-4
5	1223	23.2	975	9	US-10-192-867-2
6	1221.5	23.2	972	10	US-09-944-807-10
7	1196	22.7	976	9	US-10-099-895-32
8	1196	22.7	976	9	US-10-192-867-4
9	1152.5	21.9	1088	9	US-09-961-403-4
10	1152.5	21.9	1089	10	US-09-769-987-2
11	1152.5	21.9	1089	10	US-09-919-497-90
12	1152.5	21.9	1089	10	US-09-866-510-2
13	1150.5	21.9	1089	9	US-09-955-363-36
14	1149.5	21.8	1089	10	US-09-866-510-10
15	1148.5	21.8	1089	10	US-09-866-510-4
16	1147.5	21.8	1089	10	US-09-866-510-8
17	1146.5	21.8	1089	10	US-09-866-510-6
18	1078	20.5	1090	10	US-09-866-510-14
19	1078	20.5	1106	9	US-09-955-363-2

20	1078	20.5	1106	10	US-09-866-510-22	Sequence 22, Appl
21	1074	20.4	1106	10	US-09-866-510-16	Sequence 16, Appl
22	1073	20.4	1106	10	US-09-866-510-20	Sequence 20, Appl
23	1072	20.4	1106	10	US-09-866-510-18	Sequence 18, Appl
24	998.5	19.0	1338	9	US-10-059-585-44	Sequence 44, Appl
25	961.5	18.3	1356	9	US-10-022-939-2	Sequence 2, Appl
26	961.5	18.3	1356	9	US-10-100-405A-2	Sequence 2, Appl
27	960.5	18.2	1356	9	US-09-969-037-7	Sequence 7, Appl
28	954.5	18.1	1298	10	US-09-982-610-33	Sequence 33, Appl
29	954.5	18.1	1363	9	US-09-375-248-2	Sequence 2, Appl
30	952.5	18.1	1368	9	US-10-105-901-34	Sequence 34, Appl
31	952	18.1	1363	9	US-09-375-248-19	Sequence 19, Appl
32	943.5	17.9	1362	9	US-10-105-901-33	Sequence 33, Appl
33	941	17.9	386	9	US-09-939-833-6	Sequence 6, Appl
34	941	17.9	386	10	US-09-939-754-6	Sequence 6, Appl
35	941	17.9	386	10	US-09-939-832-6	Sequence 6, Appl
36	937.5	17.8	1367	10	US-09-766-678-2	Sequence 2, Appl
37	936.5	17.8	1363	9	US-10-105-901-32	Sequence 32, Appl
38	926.5	17.6	1367	10	US-09-919-408-6	Sequence 6, Appl
39	926.5	17.6	1367	10	US-09-872-136-6	Sequence 6, Appl
40	862	16.4	367	9	US-09-939-833-9	Sequence 9, Appl
41	862	16.4	367	10	US-09-939-754-9	Sequence 9, Appl
42	862	16.4	367	10	US-09-939-832-9	Sequence 9, Appl
43	856	16.3	367	9	US-09-939-833-12	Sequence 12, Appl
44	856	16.3	367	10	US-09-939-754-12	Sequence 12, Appl
45	856	16.3	367	10	US-09-939-832-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1

US-09-919-408-2

; Sequence 2, Application US/09919408

; Patent No. US20020072077A1

; GENERAL INFORMATION:

APPLICANT: Lemischka, Ihor R.

TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

RECEPTORS AND THEIR LIGANDS

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Imclone Systems Incorporated

STREET: 180 Varick Street

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10014

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0., Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/919,408

FILING DATE: 31-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/977,451

FILING DATE: <Unknown>

APPLICATION NUMBER: US 07/906,397

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: US PCT/US92/05401

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: TW 81102961

FILING DATE: 15-APR-1992

APPLICATION NUMBER: US PCT/US92/02750

FILING DATE: 02-APR-1992

APPLICATION NUMBER: US 07/813,593

FILING DATE: 24-DEC-1991

APPLICATION NUMBER: US 07/793,065

FILING DATE: 15-NOV-1991

APPLICATION NUMBER: US 07/728,913

FILING DATE: 28-JUN-1991

APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 992 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-919-408-2

Query Match 100.0%; Score 5264; DB 10; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALAQRDRRLLLVLSVMILETVTNDLPVIRKVLISHENNGSSAGKPSRYMRVRS 60  
Db 1 MRALAQRDRRLLLVLSVMILETVTNDLPVIRKVLISHENNGSSAGKPSRYMRVRS 60

Qy 61 PEDLOCTPRRQSEGVYEAATVEAEGSITLQVQLATPGDLSCLVWFKHSLGCOHPED 120  
Db 61 PEDLOCTPRRQSEGVYEAATVEAEGSITLQVQLATPGDLSCLVWFKHSLGCOHPED 120

Qy 121 LQNRGIVSMALLNVETQAGEYLLHQSRANVTYLVTVNVDRDTOLYVLRFRYFRKMENQ 180  
Db 121 LQNRGIVSMALLNVETQAGEYLLHQSRANVTYLVTVNVDRDTOLYVLRFRYFRKMENQ 180

Qy 181 DALLCISGEVPEPTVWVLCSSHRESCKEEGPAVVRKEEVLHFGTDIRCCARNALGR 240  
Db 181 DALLCISGEVPEPTVWVLCSSHRESCKEEGPAVVRKEEVLHFGTDIRCCARNALGR 240

Qy 241 ECTKLTIDLNOAPOSTLPQLFKVGEPLWIRCKAIHVNHGGLTWELDKALEGSYFE 300  
Db 241 ECTKLTIDLNOAPOSTLPQLFKVGEPLWIRCKAIHVNHGGLTWELDKALEGSYFE 300

Qy 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSOSALVTILEKGFINATSSOEY 360  
Db 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSOSALVTILEKGFINATSSOEY 360

Qy 361 EIDPYEKFCSVRFKAYPRIRCTWIFSQAFFCEQRLGEGYSISKFDHKNKPGYIFY 420  
Db 361 EIDPYEKFCSVRFKAYPRIRCTWIFSQAFFCEQRLGEGYSISKFDHKNKPGYIFY 420

Qy 421 AENDDAQFTKMTLNIRKPKQVLANASASQSSDGYPLPSWTWKCKSDKSPNCTEIP 480  
Db 421 AENDDAQFTKMTLNIRKPKQVLANASASQSSDGYPLPSWTWKCKSDKSPNCTEIP 480

Qy 481 EGVWNNKANRKYFGQWVSSTLNMSSAGKGLLVKCCAYNSMGTSCETIFLNSPPPPFTQ 540  
Db 481 EGVWNNKANRKYFGQWVSSTLNMSSAGKGLLVKCCAYNSMGTSCETIFLNSPPPPFTQ 540

Qy 541 DNISFYATIGLCLPFTVILVILCHYKQFRYESQLOMIOVTGPDLDNEYFYVDFRDY 600  
Db 541 DNISFYATIGLCLPFTVILVILCHYKQFRYESQLOMIOVTGPDLDNEYFYVDFRDY 600

Qy 601 DLKWEPPRENLEFGKVLGSAFGVRNATAYGISKTVGSIVQAVKMLKEKADSCKEALM 660  
Db 601 DLKWEPPRENLEFGKVLGSAFGVRNATAYGISKTVGSIVQAVKMLKEKADSCKEALM 660

Qy 661 SELKMTHTLGHHDNINVLGACTLSGPVYLIFEYCCYGDLLNVLRSKREKFRHTWEIFK 720  
Db 661 SELKMTHTLGHHDNINVLGACTLSGPVYLIFEYCCYGDLLNVLRSKREKFRHTWEIFK 720

Qy 721 EHNFSYPTFOAHSSNMPGSRVQLHPDLQDSGFNGNSIHSEDEIYENQKRLAESEE 780  
Db 721 EHNFSYPTFOAHSSNMPGSRVQLHPDLQDSGFNGNSIHSEDEIYENQKRLAESEE 780

Qy 781 EDNLNLTPEDLCCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVXICDFGLARDIL 840  
Db 781 EDNLNLTPEDLCCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVXICDFGLARDIL 840

Qy 841 SDSSVYVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPGIPVDAN 900  
Db 841 SDSSVYVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPGIPVDAN 900

Qy 901 FYKLIOSGFKMEQPEYATEGIYVMSQSWAFDSRKRSPNLTSLGCGQLAEAEACIRT 960  
Db 901 FYKLIOSGFKMEQPEYATEGIYVMSQSWAFDSRKRSPNLTSLGCGQLAEAEACIRT 960

Qy 961 SIHLPKQAAPQORGGLRAQSPQGVKIHRRS 992  
Db 961 SIHLPKQAAPQORGGLRAQSPQGVKIHRRS 992

RESULT 2  
US-09-872-136-2  
Sequence 2, Application US/09872136  
Patent No. US20020119545A1  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESS: ImClone Systems Incorporated  
STREET: 180 Varlick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/872,136  
FILING DATE: 01-Jun-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/208,786  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/09/021,324  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/07/977,451  
FILING DATE: 1992-11-19  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 992 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-872-136-2

Query Match 100.0%; Score 5264; DB 10; Length 992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MRALAQRSDRLLLVLSVMIETVTDLPVVKVLSHSHENNGSSAGKPSRYMRVGRS 60  
Db 1 MRALAQRSDRLLLVLSVMIETVTDLPVVKVLSHSHENNGSSAGKPSRYMRVGRS 60

Qy 61 PEDLQCTPRROSEGTVYEAATVEAEGSITLQVOLATPGDLSCLVWFKHSLGCGPHFD 120  
Db 61 PEDLQCTPRROSEGTVYEAATVEAEGSITLQVOLATPGDLSCLVWFKHSLGCGPHFD 120

Qy 121 LQNRGIYSMAILNVTTQAGEYLLHIQSERANYTVLFTVNVNRDTOLYVLRPFYFRKMENQ 180  
Db 121 LQNRGIYSMAILNVTTQAGEYLLHIQSERANYTVLFTVNVNRDTOLYVLRPFYFRKMENQ 180

Qy 181 DALLCISGEVPEPTVEWVLCSSHRESCKEKGPAVVRKEEVHLHFGTDIRCCARNALGR 240  
Db 181 DALLCISGEVPEPTVEWVLCSSHRESCKEKGPAVVRKEEVHLHFGTDIRCCARNALGR 240

Qy 241 ECTKLFTIDLQAQOSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWEEDKALEGSGYFE 300  
Db 241 ECTKLFTIDLQAQOSTLPQLFKVGEPLWIRCKAIHVNHGFGLTWEEDKALEGSGYFE 300

Qy 301 MSTYSTNTRMIRILLAFVSSVGRNDTGYTCTSSSKHPQSOSALVTILEKGINATSSQEEY 360  
Db 301 MSTYSTNTRMIRILLAFVSSVGRNDTGYTCTSSSKHPQSOSALVTILEKGINATSSQEEY 360

Qy 361 EIDPYEKFCFSVRKAYPRIRCTWIFSOASPCGEORGLEDGYSISKFDHKNKPGYIFY 420  
Db 361 EIDPYEKFCFSVRKAYPRIRCTWIFSOASPCGEORGLEDGYSISKFDHKNKPGYIFY 420

Qy 421 AENDDAQFTKFTNIRKKPOVLANASQASCSDDGYPLPSWTWKKSDKSPNCTEETIP 480  
Db 421 AENDDAQFTKFTNIRKKPOVLANASQASCSDDGYPLPSWTWKKSDKSPNCTEETIP 480

Qy 481 EGVWNNKANRVFGOWSSSTLNMSKAGLLVKKCCAYNSMGTSCETIFLNSPGPPFTIQ 540  
Db 481 EGVWNNKANRVFGOWSSSTLNMSKAGLLVKKCCAYNSMGTSCETIFLNSPGPPFTIQ 540

Qy 541 DNISFYATIGLCLPIVVLVILCHIKYKQPRYESOLQMIQVTGLDNEYFYVDFRDYFY 600  
Db 541 DNISFYATIGLCLPIVVLVILCHIKYKQPRYESOLQMIQVTGLDNEYFYVDFRDYFY 600

Qy 601 DLKWEFFPRENLEFGKVLGSGAFGRVWNNATAYGISKTGYSIQVAVKMLKEKADSCKEALM 660  
Db 601 DLKWEFFPRENLEFGKVLGSGAFGRVWNNATAYGISKTGYSIQVAVKMLKEKADSCKEALM 660

Qy 661 SELKMWTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLSRKREKFRHTWTETIFK 720  
Db 661 SELKMWTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLSRKREKFRHTWTETIFK 720

Qy 721 EHNFSYTFQAHNSNMPGSRREVOLHPPDLQSLGFGNSIHSDEIEYENOKRLAESEE 780  
Db 721 EHNFSYTFQAHNSNMPGSRREVOLHPPDLQSLGFGNSIHSDEIEYENOKRLAESEE 780

Qy 781 EDLNVLTEDLLCFAYQAKGNFELEFSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840  
Db 781 EDLNVLTEDLLCFAYQAKGNFELEFSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840

Qy 841 SDSSVYVGRNARLPVKWMAPESLFEGITIKSDVNSYIGILLWEIFSLGVNYPGIPVDAN 900  
Db 841 SDSSVYVGRNARLPVKWMAPESLFEGITIKSDVNSYIGILLWEIFSLGVNYPGIPVDAN 900

Qy 901 FYKLIQSGFKMEQPEYATYEGYFVMSQWAFDSRKRPSFNLTSFLGCOLAEAEACIRT 960  
Db 901 FYKLIQSGFKMEQPEYATYEGYFVMSQWAFDSRKRPSFNLTSFLGCOLAEAEACIRT 960

Qy 961 SIHLPKQAAPQORGGRLAQSPOQVKIHRERS 992  
Db 961 SIHLPKQAAPQORGGRLAQSPOQVKIHRERS 992

RESULT 3  
US-09-919-408-4  
; Sequence 4, Application US/09919408  
; Patent No. US20020072077A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/919,408  
; FILING DATE: 31-Jul-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/977,451  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 28-JUN-1992  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; APPLICATION NUMBER: US 07/679,666  
; FILING DATE: 02-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: LEM-3-7P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405  
; TELEFAX: 212-645-2054  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 993 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-919-408-4

Query Match 84.1%; Score 4429.5; DB 10; Length 993;  
Best Local Similarity 84.1%; Pred. No. 2,3e-297;  
Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

QY 1 MRALQSRDRRLLLVLSVMIETVITNODLPVIRKVLISHENNGSSAGKPSRYMRVRS 60  
DB 1 MPALA-RDAGTVPLLVVFSAMIFGITNODLPVIRKVLINHNNDSSVCKSSYPMVSES 59  
QY 61 PEDLOCTPRQSEGVYEAATVEAEGSITLQVLATPGDLSCLWVFKHSLGCPHF 120  
DB 60 PEDLGALRQSSGVYEAADVEDVSITLQVLVDPAGNISCLWVFKHSLNCPHF 119  
QY 121 LQNRGIVSMAILNVETQAGEYLLHTQSRANYTVLFTVNVDRDTOLYVLRPFYFRKMEQ 180  
DB 120 LQNRGVSMVILKMTQAGEYLLFQSRATNTILFTVIRNTLYTLRRPFYFRKMEQ 179  
QY 181 DALLCISEGVPEPTVWVLCSSHRECKEKGPAVVRKEEVHLFGTDIRCCARNALGR 240  
DB 180 DALVCISEVPEPIVWVLCDSOGESKEESPAVVRKEEVHLFGTDIRCCARNELGR 239  
QY 241 ECTKLTIDLNOPOSTLPQLFKVGEPLWIRCKAIHVNHGGLTWELDKALEEGSYFE 300  
DB 240 ECTRLFTIDLNOTPOTLPQLFKVGEPLWIRCKAVHVNHGGLTWELDKALEEGSYFE 299  
QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSOSALVTILEKGFINATSSOEY 360  
DB 300 MSTYSTNRTMIRILLAFVSSVARNDTGYTCSSSKHPSOSALVTIVGKGFINATSSSEY 359  
QY 361 EIDPYEKFCSVRFKAYPRICRTWTFPSQASFPCEQRLGLEDGYISIKFCDHKNKPGYIEY 420  
DB 360 EIDQYEEFCFSVRFKAYPOICRTWTFPSRKSFPCEQKGLDNGYISIKFCNKHQPGYIEH 419  
QY 421 AENDDAQFTKMTLNIRKPKOVLANASASQSCSDGYPLPSTWTKKCDKSPNCTEETP 480  
DB 420 AENDDAQFTKMTLNIRRRKQVLAESAQSCSDGYPLPSTWTKKCDKSPNCTEET 479  
QY 481 EGVNKKANKRVFGQWVSSTLNMSAGLLVKCCAYNSMTGSCETIFLNSGPPFPFIQ 540  
DB 480 EGVNKKANKRVFGQWVSSTLNMSAIGLLVKCCAYNSLGTSCETIFLNSGPPFPFIQ 539  
QY 541 DNISFYATIGLCLPFTVILVILICHYKKQFRYESQLQMIQVTPGLDNEYFYVDFRDY 600  
DB 540 DNISFYATIGVCLLFTVLTLICHYKKQFRYESQLQVQVVTGSSDNEYFYVDFREY 599  
QY 601 DLKWEPPRENLEFGKVLGSGAGRVNATAYGISKTVGSIOQAVKMLKEKADSCREALM 660  
DB 600 DLKWEPPRENLEFGKVLGSGAGRVNATAYGISKTVGSIOQAVKMLKEKADSCREALM 659  
QY 661 SELKMTHTLGHNDINVLGACTLSPGVYLIFCYCCYGLDNLVLRKREKFRHTWTEIFK 720  
DB 660 SELKMTHTLGHNDINVLGACTLSPGVYLIFCYCCYGLDNLVLRKREKFRHTWTEIFK 719  
QY 721 EHNFSYPTFOAHNSMPSGSEVQLHPDLQSLGFGNGNSIHSEDETEYENQKRLAEFE 780  
DB 720 EHNFSYPTFOAHNSMPSGSEVQLHPDSDQISGLHNSFHSEDETEYENQKRL--EE 777  
QY 781 EDNLVTFEDLLCFAYQVAKMEFEFKSCVHRDLAARNVLTGHVKVVKICDFGLARDIL 840  
DB 778 EDNLVTFEDLLCFAYQVAKMEFEFKSCVHRDLAARNVLTGHVKVVKICDFGLARDIM 837  
QY 841 SDSYVVRGNARLPVKWMPESLFEGLIYTKSDVWSVIGILLWIFSLGVNYPGIPVDAN 900  
DB 838 SDSYVVRGNARLPVKWMPESLFEGLIYTKSDVWSVIGILLWIFSLGVNYPGIPVDAN 897  
QY 901 FYKLIOGFKMEOPFVATEGIYFVMOSWAFDSKRPSPFNLTSLFLGCOLAEAEAC-- 957  
DB 898 FYKLIOGFKMEOPFVATEGIYIIMOSWAFDSKRPSPFNLTSLFLGCOLAEAEAMQN 957  
QY 958 ----IRTSIHLPKAAPQQRG--GLRAQSPQOVK 986  
DB 958 VDCRVSECPHYQNRPFPSREMDLGLLSPQAQVE 991

## RESULT 4

US-09-872-136-4

; Sequence 4, Application us/09872136

; Patent No. US20020119545A1

GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOIROPOTENT HEMATOPOIETIC STEM CELL RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESS: ImClone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/872,136  
FILING DATE: 01-Jun-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/208,786  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/09/021,324  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/07/977,451  
FILING DATE: 1992-11-19  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: TW 81102961  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-872-136-4

Query Match 84.1%; Score 4429.5; DB 10; Length 993;  
Best Local Similarity 84.1%; Pred. No. 2.3e-297;  
Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;  
QY 1 MRALQSRDRRLLLVLSVMIETVITNODLPVIRKVLISHENNGSSAGKPSRYMRVRS 60  
DB 1 MPALA-RDAGTVPLLVVFSAMIFGITNODLPVIRKVLINHNNDSSVCKSSYPMVSES 59  
QY 61 PEDLOCTPRQSEGVYEAATVEAEGSITLQVLATPGDLSCLWVFKHSLGCPHF 120  
DB 60 PEDLGALRQSSGVYEAADVEDVSITLQVLVDPAGNISCLWVFKHSLNCPHF 119  
QY 121 LQNRGIVSMAILNVETQAGEYLLHTQSRANYTVLFTVNVDRDTOLYVLRPFYFRKMEQ 180



```
Db 120 LQNRGVSMVLKMTQAGEYLLFQSEATNYTLFTVSIKNTLLYTLRRPYFRKMEQ 179
Qy 181 DALLCISEGVPEPTVEWVLCSSSHRECKEKGPAVVRKEEKLHELFGTDIRCCARNALGR 240
Db 180 DALVCISEVPPEVIEWVLCDSOGCKEESPAVVRKEEKLHELFGTDIRCCARNELGR 239
Qy 241 ECTKLFTIDLNOAPOSTLPQLFKYGEPLWIRCKALHVNHRGFLTWELKEDKALEEGSYFE 300
Db 240 ECTRLFTIDLNOTPQTLPLQFKYGEPLWIRCKAVHVNHRGFLTWELKEDKALEEGSYFE 299
Qy 301 MTSYTNTRMIRILLAFVSSVGRNDTGYTSCSSKHPSQSALVTILEKFINATSSQEBY 360
Db 300 MTSYTNTRMIRILLAFVSSVGRNDTGYTSCSSKHPSQSALVTIVGKFINATSSQEBY 359
Qy 361 EIDPYEKCFCFVRKAYPRICRTWTFISQAFCEQGLDGYISKFCDHKNKPGYIY 420
Db 360 EIDQYEEFCFVRKAYPOICRTWTFISQAFCEQGLDGYISKFCDHKNKPGYIY 419
Qy 421 AENDDAQFTKMTLIRKRPQVLANASASOASCSGDYPLPSWTWKCKSDKSPNCTEIP 480
Db 420 AENDDAQFTKMTLIRKRPQVLANASASOASCSGDYPLPSWTWKCKSDKSPNCTEIT 479
Qy 481 EGVNKKANKRVFGQWVSSTLNMSEAGKGLLVKCCAYNSMGTSCTETIFLNSPGPFPIQ 540
Db 480 EGVNKKANKRVFGQWVSSTLNMSEAIKGLVLKCCAYNSMGTSCTETILLNSPGPFPIQ 539
Qy 541 DNISYATIGLCLPIVVLVILCHIKYKQFRESQLOMQVITGPDNDNEYFYVDFRDEY 600
Db 540 DNISYATIGVCLLPIVVLVILCHIKYKQFRESQLOMQVITGPDNDNEYFYVDFRDEY 599
Qy 601 DLKWEFFPRENLEFGLVGSAGFGRVNNATAYGISKTGYSIQVAVKMLKADCKEALM 660
Db 600 DLKWEFFPRENLEFGLVGSAGFGRVNNATAYGISKTGYSIQVAVKMLKADCKEALM 659
Qy 661 SELKMTHLGHNDINVLNLLGACTLSGPVLIFFEYCCYGDLLNLYRSKREKFRHTWTEIFK 720
Db 660 SELKMTQLGSHENIVNLLGACTLSGPVLIFFEYCCYGDLLNLYRSKREKFRHTWTEIFK 719
Qy 721 EHNFSYPTFOAHNSMPSGSRVQLHPLDOLSGFNGSIHSEDEIEYENOKRLAEDEE 780
Db 720 EHNFSYPTFOAHNSMPSGSRVQLHPLDOLSGFNGSIHSEDEIEYENOKRLAEDEE 777
Qy 781 EDNLVTFEDILCFAYQVAKGMEFLKFCVHRDLAARNVLTGKVKVVICDFGLARDIL 840
Db 778 EDNLVTFEDILCFAYQVAKGMEFLKFCVHRDLAARNVLTGKVKVVICDFGLARDIN 837
Qy 841 SDSYVVRGNARLPVKWMAPESLFEGITIKSDVNSYGILLWEIFSLGVNYPGIPVDAN 900
Db 838 SDSYVVRGNARLPVKWMAPESLFEGITIKSDVNSYGILLWEIFSLGVNYPGIPVDAN 897
Qy 901 FYKLQSGFKMEOPPYATEGYIFVMSQWAFDSRKRPFNPNTSLGCOLAEAEAC 957
Db 898 FYKLQSGFKMDQPYATEGYIFVMSQWAFDSRKRPFNPNTSLGCOLAEAEAMYN 957
Qy 958 ---IRTSIHLPKQAAPOQR-GLRAQSPOROVK 986
Db 958 VDRVSECPHYVONRRPFSREMDLGLLSPQAQVE 991
```

## RESULT 5

```
US-10-192-867-2
; Sequence 2, Application US/10192867
; Publication No. US2003008466A1
; GENERAL INFORMATION:
; APPLICANT: BLUME-JENSEN, Peter
; TITLE OF INVENTION: SYSTEM AND METHOD FOR CONTROLLING MALE FERTILITY
; FILE REFERENCE: SALKINS.002C1
; CURRENT APPLICATION NUMBER: US/10/192.867
; PRIOR APPLICATION NUMBER: 2002-07-08
; PRIOR APPLICATION NUMBER: US 60/175,625
; 2000-01-11
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00573
; 2001-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 975
; TYPE: PRN
; ORGANISM: MUS MUSCULUS
US-10-192-867-2

Query Match      23.2%; Score 1223; DB 9; Length 975;
Best Local Similarity 32.1%; Pred. No. 4.3e-76;
Matches 324; Conservative 172; Mismatches 331; Indels 182; Gaps 39;

Qy 36 CVLISHENNGSAGKPSYRMVRGSPEDLOCTPRRSEGTVYEAAATVEVEAESGITLQVQ 95
Db 12 CVLLVLLRGQTATSPSA-----SPGEPSPPSIHPAQSELIVEAGD-TL--- 54
Qy 96 LATPGDLSCL-----WFEKHSGLGQPHF-----DLQNRGIVSMALLNVTETQAGEYLL 144
Db 55 -----SLTCIDPDFFVRWTEK-----TYFNEMVENKKNEWIOEKA-----EATRTGYTC 98
Qy 145 HIOSEANVTYVLTVMVRD-TQLYVLRPYFRKMNODALL-----CIS 187
Db 99 ---SNSNGLTSSIIYFVRDPAKLFVLGLPLFGK-EDSDALVRCPLTDPQVSNYSLEICDG 154
Qy 188 EG-----VPPTVEWVLCSSHRECKEKGPAVVRKEEKLHELFGTDIRCCARNALGR 240
Db 155 KSLPTDLTVPNP-----KAGITIKNVKRAYH-----RLCVRCAAOR 191
Qy 241 ECTKL-----FTIDLNOA-----FLKVGCEPLWIRCKALHVNHRGFLTW--- 286
Db 192 DGTWLSHDAFTLKVRRAIKAIPIVSVPEFTHLLKGGDTFTVVTCTIKDVSTSVSNMWMKMN 251
Qy 287 -----ELEDKALEEGSYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYTSCSSKHPSQ 339
Db 252 POQHIAQVKHNSHRGDF-----NYERQETLT-----ISSARVDDSGVFEMCVANNTFGS 301
Qy 340 SALVT---ILEKFINATS-SQEEYIDIDYEKFCFVRKAYPR-IRCTWIFSOASPFCE 394
Db 302 ANVTTLKVVKEGFINISPVKNVTTFVTGENDVLDVVEYEAYPKPEHQOVIYNRT--SA 359
Qy 395 QRGE-----DGVYSISKCDH-----KNKPEYIFVENDDAQFTKMTLIRKRPQVL 443
Db 360 NKGKDYKSDKNSIRYVQNLRLTLKGTGGTYTFLVNSDSASASVTFVYVNTKREIL 419
Qy 444 A---NASASOASCSGDYPLPSWTWKCKDKSPNCTBEI-PEGVNNKANKRVFGQWVSS 500
Db 420 TYDRLINGMLQCVAEQFPEPTIDWYFCTGAEQRCTTPVSPDVQVQVNSVSPGKLVQS 479
Qy 501 TLNMSEAGKGLLVKCCAYNSMGTSCTETIFLNSPGPFPFQDNISFYATIGLCLPFIV--- 557
Db 480 SIDSSVFRHNGTVCKEASNDVGKS--SAFFN-----FAFKEIQOAHLTFTPLLLGFVVAAG 533
Qy 558 ---VLITVLVICHYKKOFRYESLOMI-OVTGPDNDVEYVDFRDEYDLKWEFPRENLEF 613
Db 534 AMGLIVVLTLYKLPMTVEVQKVVVEEING---NNVYIDTQLPDYDHKWEPRNRLSF 590
Qy 614 GKVLGSGAFGRVNNATAYGISKTGYSIQVAVKMLKADCKEALMSELKMTHLGHHD 673
Db 591 GKTGLGAGAFKGVVEATAYGLIKSDAAMTAVAKMLKPSAHLTEREALMSELKVLVSYLGNHM 650
Qy 674 NTVNLGACTLSGPVLIFFEYCCYGDLLNLYRSKREKPF-----HRTWTEIFKEHNPSSY 727
Db 651 NTVNLGACTVGGTTLVITEYCCYGDLLNLFRRKRDSFIFSKQEEQAEALYKNLLHSTE 710
Qy 728 PTFQAHSSSM---PGSREVQLHPPLDOLSGFNGSIHSEDEIEYENOKRLAEDEEDLN 784
Db 711 PSCDS--SNEYDMKKPGVSIV-VPTKTDK-----RRSARIDSYIRDVDTPAIMDEDLALD 763
Qy 785 VLTFFEDLLCFAYQVAKGMEFLKFCVHRDLAARNVLTGKVKVVICDFGLARDILSDSS 844
Db 764 L---DDLFSFQVAKAMAFKASKNCIHRDLAARNLLTHGRITKICDFGLARDIRDSN 820
```

```

QY      845  YVYRGHARLPVKWMADESLEFEGIYTIKSDVMSYGILLWEIFSLGVNYPYCGIPYDANFYKL 904
      |||:|||||:|||||:  ||:|||||:  |||:||||:  |||:||||:  |||:
Db      821  YVYVKGHARLPVKWMADESIFSCVYTFESDVMSYGIFLWELFSLGSSPYPGMPVDSKEYKM 880

QY      905  IQSGFKMEQPFYATEGIYFMOSCHAFDSKRKPSFNLTSFLGCCQAAEA 953
      |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:  |||:
Db      881  IKGFRFVMSPEHAEMYDMVKTCWDADPLKRPTFKQVVLIEKQIIDS 929

RESULT 6
US-09-944-807-10
; Sequence 10, Application US/09944807
; Patent No. US2002011949A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim Pharma KG
; TITLE OF INVENTION: Method for identifying substances which positively
; TITLE OF INVENTION: influence inflammatory conditions of chronic
; TITLE OF INVENTION: inflammation of airway diseases
; FILE REFERENCE: 082_00n
; CURRENT APPLICATION NUMBER: US/09/944,807
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: UK 0021484.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-944-807-10

```

Query Match	23.2%	Score 1221.5;	DB 10;	Length 972;
Best Local Similarity	32.5%;	Pred. No. 5.4e-76;		
Matches	330; Conservative 152;	Mismatches 331;	Indels 201;	Gaps 33;

  

Qy	76	VYEAATVEVAESGSITLQVOLATPGDLSCLWVKFHSLSGCPHFDFLONRGIVSMAIL---	132
Db	22	VIEPSVELVVKPGAT--VTLRGVNGSVGEWDGPPS-----PHTLVISDG--SSSILSTN	72
Qy	133	NVETQAGEYLLHIQSERANYTVLTINVNRDTQLYLRRPY-----FRKMNODALL-C	185
Db	73	NATEONTGTCTCEPDPLGGSAAIHLVXKP-----ARPNVLAEVVFEVDQDALLPC	127
Qy	186	ISEGVPEPTVENWLCSHURESCKEEGPAVVRKEK-----VLHE---LF	226
Db	128	L---LTDPVL-----EAGVSLVRGRPLMRHTNYSFSPWHGFTHIRAKFIQ	171
Qy	227	GTDTRCCARNALRECTKLFTIDLNQAQPSTLP-----OLFLKVGGEPLMIRCKAI	276
Db	172	SDYQCSALMG-GR---KVMSISIFLKVKVIPGPPALTLPaelvRlRGEAAQIVCSAS	227
Qy	277	HVNHFGTLWEDEKALBEGSYFMSTYSTNRTMIIRILLAVSSVGRNDTGYTCSS--	334
Db	228	SDVNFDVFLOHNNTKL---AIPQSDPHNNRYQ-KVLTNLNDQDVOFHAGNYSVCVASNV	283
Qy	335	--KHPSOASLYTILEKGFINATSSOE-EYEIDPYEKFCFSRFKAPRIR-CTWIFSQAS	390
Db	284	QKHK-STSNMFFRVVESAYLNSUSSEONLQIOWTVGBGLNKVMVEAYPGLOGFNWTY----	338
Qy	391	FPCEORGLEDSYISKFCOHKNKP-----GEYIFYAEN	423
Db	339	-----LGPFSDHOPEPKLANATTKTDTYRHFTTLLSLPLKPSAGRVSFLARN	385
Qy	424	DDAQTKMTLINRKPKOV----LANASQASCSSDCYPPLPSTWTKKCDKSNCPTEE	478
Db	386	PGWRALAFELTLRYPPPEVSVIWTEPIINGSQT--LLCAASGYQPQNVTWLQCSGHTDRCDEA	444
Qy	479	IPEGWN---KKANRKVFEGQWVSSTLNMSEPAGLLVKCCAYNSMGTSCETIFLNSPG	534
Db	445	QVLQWDDPYEPVLUSEGPEPHKVTYVOSLLTVELEHNQHYECRAHNSVGGSWAFPIBSAG	504
Qy	535	PFPFIQDNISFYATIGLC---LPFITVLIIVLICHHKYKKQFYESOLOMQIQTGPLDNIYF	591

166 PKAGIMIKSVKRAYHRLCLHCSVDQEG-----KSVLSEKFIKVRPAFK-----A 210  
245 LFTIDLNAQPOSTLPOLFLKVGEPDLWIRCKAIHVNHGFLTWLEDEK--ALEE----- 295  
211 VPVSVSKA-----SYLLREGEETVTCTIKDVSSSVYTWKRENSQTKLOEKYNSWHH 264  
296 GSYFEMSTYNTMIRILLAFVSSVGRNDTGYTSCSSKHPQSALVTILE---KGFIN 352  
265 GDF-----NYERQATLT-----ISSARVNDGVPFCYANNTFGSANVTTLVVDKGFIN 314  
353 A-TSSQBEYEIDPYEKFCSVRFKAYPR-IRCTWIFSOASF--PCEQGLEDDGYISKFC 408  
315 IFPMINTTVFVNDGENDLIVEYEAFFPKPEHQOYIMNRTFTDKWEDYPKSESNIRYV 374  
409 DHKN-----KPEYIFYAENDDAQFTKMTFLNIRKQPOLA--NASASQSCSSDGY 459  
375 SELHLTRKLTGEGTGYTFLVNSDVNAIAFNVYNTKPEILTYDRLVNGMLQCAAGFP 434  
460 LPSWTWKKCDKSPNCTEEI-PEGVWNNKANRKFVGOWSSSTLNNSEAGKGLLVKCCAY 518  
435 EPIDWYFCGTEQRCASVLPVDVOTLNSGPPFGKLVQSSIDSSAFKNGHNGVECKAY 494  
519 NSMGTSCETIFLNSPGPPFTQDN-----ISFYATIGLCLPFIIVLVILIC 564  
495 NDVGKT--SAYFN---FAFGKNNKEQIHPHTLFTPLLLIGFVIVAGM---MCIIVMILT 544  
565 HKYKKQFRIESQLOMI-QVTGPLDNEYFYVDFRDYEDLKWEPFRENLEFGKVLGSGAGF 623  
545 YKYLQKPMYEQWKVVEEING---NNVYIDPTOLPYDTHKWEFPNRLSFGKTLGAGAF 601  
624 RVNMTAYGTSKTVSIOVAVKMLKEKADSCKEALMSELKMMTHLGHHDNIVNLGACT 683  
602 KVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLSTLGNHMNIVNLGACT 661  
684 LSGPVYLIFCYCCYGDLLNLRKREKFRHTWTETEIFKEHNFSSYPTTQAHNSNSMPGSR 743  
662 IGGPTLVITEYCCYGDLLNLRKRRKDFICSQEDHAEALYKNLLHSKSSCSDSTNEY 721  
744 VOLHPPDLQOL---SGFNNGSHSEDEIYENOKRLAEEDLNLVTFEDLLCFAYQVAK 800  
722 MDMPGVSVYVPTKADRRSVRIGSYIERDVTPTAIMEDELALDL---EDLLSFSYQVAK 778  
801 GMEFLEPKSVHRDLAARNLVTHGKVVVKICDFGLARDILSDSSYVVRGNARLPVKWAP 860  
779 GMAFLASKNCIHRDLAARNILLTHGRITKICDFGLARDIKNDNSYVVRGNARLPVKWAP 838  
861 ESLFEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPYDANFYKLIQSGFKMEQPFYATEG 920  
839 ESIFNCVYTFESDVSWSYGIFLWELFSLGSSPYGMPVDSKFKYKMKIEGFRMLSPHAPAE 898  
921 IYFVWQSCWAFDSKRSPFNLTSLGCOLAEB 953  
899 MYDIMKTCWDADPLKRPTRFKQIVOLIEKOISES 931

RESULT 8

US-10-192-867-4  
; Sequence 4, Application US/10192867  
; Publication No. US2003008466A1  
; GENERAL INFORMATION:  
; APPLICANT: BLUME-JENSEN, Peter  
; APPLICANT: HUNTER, Tony  
; TITLE OF INVENTION: SYSTEM AND METHOD FOR CONTROLLING MALE FERTILITY  
; FILE REFERENCE: SALKINS.002C1  
; CURRENT APPLICATION NUMBER: US/10/192,867  
; CURRENT FILING DATE: 2002-07-08  
; PRIOR APPLICATION NUMBER: US 60/175,625  
; 2000-01-11  
; PRIOR APPLICATION NUMBER: PCT/US01/00573  
; 2001-01-08  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4

LENGTH: 976  
TYPE: PRT  
ORGANISM: HOMO SAPIENS  
US-10-192-867-4  
Query Match 22.7%; Score 1196; DB 9; Length 976;  
Best Local Similarity 31.4%; Pred. No. 3.1e-74;  
Matches 312; Conservative 172; Mismatches 361; Indels 148; Gaps 32;  
QY 36 CVLISHENNGSAGKPPSSRYMVRGSPEDLOCTPRQSEGTVYEAATVEVAESGSTITLQVO 95  
DB 12 CVLLLLLRVQSGSSOPS-----VSPGEPSPSIHPGKSDLIVRGD-EIRLL 57  
QY 96 LATPGDLSCLVAFKHSLSGCQPHFDLQNRGIVSMALLNVTETQAGEYLLHIOSEANITYV 155  
DB 58 CTDPGFVK--WTFE---ILDETENKQNEWITEKA-----EATNTKTYCTNKHGUSNIIY 108  
QY 156 LFTVNRVD-TOLYVLRPRYFRKMEQDALLCI-----SEG-----VPE 192  
DB 109 VF---VRDPAKFLVDRSLYKEDNDTLVRCPLTDPEVTNYSLKGCQKPLPKDLRFIPD 165  
QY 193 PTVEWVLCSSHRE-----SCKEEGPAVVRKEBEKVLHELFGTDIRCCARNALGRECTK 244  
DB 166 PKAGIMIKSVKRAYHRLCLHCSVDQEG-----KSVLSEKFIKVRPAFK-----A 210  
QY 245 LFTIDLNAQPOSTLPOLFLKVGEPDLWIRCKAIHVNHGFLTWLEDEK--ALEE----- 295  
DB 211 VPVSVSKA-----SYLLREGEETVTCTIKDVSSSVYTWKRENSQTKLOEKYNSWHH 264  
QY 296 GSYFEMSTYNTMIRILLAFVSSVGRNDTGYTSCSSKHPQSALVTILE---KGFIN 352  
DB 265 GDF-----NYERQATLT-----ISSARVNDGVPFCYANNTFGSANVTTLVVDKGFIN 314  
QY 353 A-TSSQBEYEIDPYEKFCSVRFKAYPR-IRCTWIFSOASF--PCEQGLEDDGYISKFC 408  
DB 315 IFPMINTTVFVNDGENDLIVEYEAFFPKPEHQOYIMNRTFTDKWEDYPKSESNIRYV 374  
QY 409 DHKN-----KPEYIFYAENDDAQFTKMTFLNIRKQPOLA--NASASQSCSSDGY 459  
DB 375 SELHLTRKLTGEGTGYTFLVNSDVNAIAFNVYNTKPEILTYDRLVNGMLQCAAGFP 434  
QY 460 LPSWTWKKCDKSPNCTEEI-PEGVWNNKANRKFVGOWSSSTLNNSEAGKGLLVKCCAY 518  
DB 435 EPIDWYFCGTEQRCASVLPVDVOTLNSGPPFGKLVQSSIDSSAFKNGHNGVECKAY 494  
QY 519 NSMGTSCETIFLNSPGPPFTQDN-----ISFYATIGLCLPFIIVLVILIC 564  
DB 495 NDVGKT--SAYFN---FAFGKNNKEQIHPHTLFTPLLLIGFVIVAGM---MCIIVMILT 544  
QY 565 HKYKKQFRIESQLOMI-QVTGPLDNEYFYVDFRDYEDLKWEPFRENLEFGKVLGSGAGF 623  
DB 545 YKYLQKPMYEQWKVVEEING---NNVYIDPTOLPYDTHKWEFPNRLSFGKTLGAGAF 601  
QY 624 RVNMTAYGTSKTVSIOVAVKMLKEKADSCKEALMSELKMMTHLGHHDNIVNLGACT 683  
DB 602 KVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSELKVLSTLGNHMNIVNLGACT 661  
QY 684 LSGPVYLIFCYCCYGDLLNLRKREKFRHTWTETEIFKEHNFSSYPTTQAHNSNSMPGSR 743  
DB 662 IGGPTLVITEYCCYGDLLNLRKRRKDFICSQEDHAEALYKNLLHSKSSCSDSTNEY 721  
QY 744 VOLHPPDLQOL---SGFNNGSHSEDEIYENOKRLAEEDLNLVTFEDLLCFAYQVAK 800  
DB 722 MDMPGVSVYVPTKADRRSVRIGSYIERDVTPTAIMEDELALDL---EDLLSFSYQVAK 778  
QY 801 GMEFLEPKSVHRDLAARNLVTHGKVVVKICDFGLARDILSDSSYVVRGNARLPVKWAP 860  
DB 779 GMAFLASKNCIHRDLAARNILLTHGRITKICDFGLARDIKNDNSYVVRGNARLPVKWAP 838  
QY 861 ESLFEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPYDANFYKLIQSGFKMEQPFYATEG 920  
DB 839 ESIFNCVYTFESDVSWSYGIFLWELFSLGSSPYGMPVDSKFKYKMKIEGFRMLSPHAPAE 898  
QY 921 IYFVWQSCWAFDSKRSPFNLTSLGCOLAEB 953  
DB 899 MYDIMKTCWDADPLKRPTRFKQIVOLIEKOISES 931







```

; LENGTH: 1089 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-955-363-36

```

Query Match	21.98;	Score 1150.5;	DB 9;	Length 1089;
Best Local Similarity	32.29;	Pred. No. 5e-71;		
Matches 313;	Conservative 154;	Mismatches 327;	Indels 177;	Gaps 32;
QY	94	VOLATPGDLSCLWFKHSSSLGQCFHDLQNRGIVSMALNVTETOAGBY--LLHIQSER	151	
DB	39	VOLNSSFSLR-----FGESEVSWQ--YPMSEBSSDVEIRN--EENSGLFTVLEVS	92	
QY	152	NYTVLFTVNVVDTOYLVLRPYFRKMENQDALLCISEG-----VPEPTV-----	197	
DB	93	AHTGLXYTCYNNHQ-----TEENE-----LEGRHIYIVPDDVAFVPLGMDY	136	
QY	198	VLCSSHRES-----CK--EEGPVAVRKEEVL-----HELFGTDIRCCARNALGR	240	
DB	137	LVIIVEDDDSAIIPCTTDPETPVTLHNSGVVPASYSRQSGFNGFTTVPPICEATV	196	
QY	241	ECTKLFTIDLNAQPOSTLPQLFL-----KVGEPLWIRKAIHVNHGFLTW-----	287	
DB	197	---KFQITPFNVYALKATSELDLEMEALKTVYKSGETIWTVC-AVFNEVVVDLQW	252	
QY	288	LEDKALEEGSYFEMSTYSTNTMIRILLAFVSSVGR---NDTGYTCSKK-----	339	
DB	253	VKGAGI-----TILEETKVPISIKLVYTLTVPATVKOSGDGYECAARQATREV	303	
QY	340	SALVTILEKEGF--INATSSQBEYIDPYEKFCSFYREKAYPRIORTWIFSOASF	397	
DB	304	KVTISVHEKGFIETKPTFSQLE-AVNLHEVKHFVVEVRAYPPIRISLWLNKLT	356	
QY	398	LEDGYSTSKFCD-----HKKK-----PEGYFYAENDDAQFTMTLNIRKRP	441	
DB	357	IENITEITDDVEKIQEIRYSRKLKLIRAKEEDSGHYTIVAQNEDAVKSYYTELL	416	
QY	442	VL-----ANASASQASCSGDGYPPLSPWTKKCDKSPNCTEETPEGVWKNK	494	
DB	417	ILDVDDHGGSTGGOTVRCATGPTLPDIEWIKCD-IKKCNNETS---WTILANN	468	
QY	495	QWYSS-----STLNNSEAGKLLVKKCCAYNSMGTSETIFLNSPGFPPTQ	540	
DB	469	--VSNIIITEHSRDRSTVEGRVTEFAKVEETAVRCLAKNLLGAENREKLVA	522	
QY	541	DNISFYATIGLCLPFIVVLIVLICHKYKQFYRESOLOMIQVTPGLDNEYFVDFR	600	
DB	523	SELTVAANVLVLVIVISLVLVVMKQPRYEIRWRVIESISPDGHEIYVDDMQL	582	
QY	601	DLKWEFFPRENLEFGKVLGSGAFGRVMNATAYGISKTGYSIQVAVKMLKEKAD	660	
DB	583	DSRWEFFPRDGLVLGRVLGSGAFGVWEGTAYGLSRQSPVMKVMKMLKPTARS	642	
QY	661	SELKMMTHLGHHDNVLNGLACTISLGPVYLIFECYCCYGDLLNLYLRSKEKF	711	
DB	643	SELKIMTHLGHPLNVLNGLACTSGPIYIITECYFGDGLVNLHKNRDSFSLSH	702	
QY	712	-----HRTWTEIFKEHNFSSYPTFOAHNSMMP--GSREVLQHPPLDQLSG	755	
DB	703	KELDIFGLNPADESTRSVSVILSFENNGDYMCKQADITQYVPMLEKEVSKY	762	
QY	756	FNGNSIHSEDIIEYENOKRLAEEREDLNVLTFDLLCFAYQAVAKGMFFLEFK	815	
DB	763	DRPASYKKKMLSDSEKVNMLSDNSEGLTLL--DLSFTYQVARGMEFLASKNC	819	
QY	816	AARNVLYTHGVKVICDFGLARDIILSDSSYVVYVGRNARLPVKWMAPESEF	875	
DB	820	AARNVLAQGIKVICDFGLARDIHSDSNYSVKGSTFLPVKWMAPESIFDNL	879	
QY	876	SYGILLWEIFSLGYNPFGPIVDANFYFKLIQSGFKMEQPFYATEGIYFVMS	935	

```

Db      880  SYGILLWEIFSGCTPYPGMMVDSTFINKISGYRMAKPDHATSEVYEIMVKCWNSEPEK  939
Qy      936  RPSFPNLTSL 946
      |||| :| :
Db      940  RPSFVHLSEIV 950

RESULT 14
US-09-866-510-10
; Sequence 10, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKONO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISE
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-510-10

```

[illegible]



QY 595 FRDYEDLKWEPPRENLEFGKVLGSGAFGRVNMATAYGISTGVSIQVAVKMLKEKADSC 654  
DB 577 PMQLPYDSRWEPDRGLVLRVLSGAGKVGVEGTAYGLSRQPVKVMKVPKMTARSS 636  
QY 655 EKEALMSELKMTWTHLGHNDINVLGACTLGGPVYLIFFEYCCYDGLNLYLRSKRKEF--- 711  
DB 637 EKOALMSELKMTWTHLGHNLINVLGACTKSGPIIITEYCFYDGLVNVYLNHKNRDSFLSH 696  
QY 712 -----HRTWTEIFEKHNFSYPTFOAHNSNMP--GSREVQLHPP 749  
DB 697 HPEKPKKELDIFGLNPADSTRSYVLSFENGDYMDMKQADTTQYVPMLEKREVSYSKD 756  
QY 750 LDQSGFNGNSIHSEDEIYENOKRLAEEDLNVLTFEDLLCFAYQVAKGMFELEFS 809  
DB 757 IORSLYDRPASYKKSMLESDSEKLNLLSDNSEGLTLL--DLSFTYQVARGMEFLASKN 813  
QY 810 CVHRDLAARNVLVTHGKVVVKICDFGLARDILSDSYVVRGNARLPVKWMAPELFEIYT 869  
DB 814 CVHRDLAARNVLLAOGKIVKICDFGLARDIMHDSNVYSGSTFLPMKWMAPESIFDNLTY 873  
QY 870 IKSDVMSYGILLWEIFSLGVNYPGIPVDANFYKLIQSGFKMEQPFYATEGIYFVMQSCW 929  
DB 874 TISDVMWSYGILLWEIFSLGTPYPCGMVDSFTYFNKIKSGYRMAKPDHATSEVYEIMVKCW 933  
QY 930 AFDSRKRPSFPNLTSL 946  
DB 934 NSEPEKRPSFYHLSEIV 950

RESULT 15  
US-09-866-510-4  
; Sequence 4, Application US/09866510  
; Patent No. US20020113041  
; GENERAL INFORMATION:  
; APPLICANT: KAZLAUSKAS, ANDRIUS  
; APPLICANT: IKUNO, YASUSHI  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES  
; CURRENT APPLICATION NUMBER: US-09/866,510  
; CURRENT FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 60/250,747  
; PRIOR FILING DATE: 2000-12-01  
; PRIOR APPLICATION NUMBER: 60/289,103  
; PRIOR FILING DATE: 2001-05-07  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 1089  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-866-510-4

Query Match 21.8%; Score 1148.5; DB 10; Length 1089;  
Best Local Similarity 32.0%; Pred. No. 6.8e-71;  
Matches 313; Conservative 155; Mismatches 320; Indels 189; Gaps 33;

QY 94 VOLATPGDLSCLWVFKHSLGCPHFDLQNRGIVSMALINVTETQAGEY--LLHIOSERA 151  
DB 39 VOLNSSFSLRC---FGSEVSWQ--YPMSEESSDYEIRN-EENNSGLFTVLEVSASA 92  
QY 152 NTVLFTVNVVDTOLVLRPRYFRKMNODALLCISEG-----VPEPTV-----EW 197  
DB 93 AHGLTTCYNNHTQ-----TEENE-----LEGRHIYYVDPDPAVFPPLGMDTY 136  
QY 198 VLCSSHRES---CK---BEGPAVRKEEKL-----HELFGTDIRCCARNALGR 240  
DB 137 LVIVEDDDSAIIPCRITDPTPTVLHNSGVWPASYSRQGFNGTFTVGYPICEATVKKG 196  
QY 241 ECTKLETDLNQAPQSTLPQLFL-----KVGPLWIRCKAIHNVHGFGLTW-----E 287  
DB 197 ---KFTQIPFNVALKATSELDLEMEALKTVYKSGETIVVTC-AVFNNNEVVDLQWTYPGE 252

Search completed: May 27, 2003, 14:45:05  
Job time : 38.4877 secs

QY 288 LEDKALEGSEFEMSTYSTNRTMIR-----LLAFVSSVGR---NDQYTYTCSSK--- 335  
DB 253 VKGKGI-----TMLBEIKVPSIKLVTLVPEATVKDSDGYDECAARQATR 297  
QY 336 --HPSQSALVTTLEKGF--INATSSQEEYEDPYEKFCEFSVREKAYPRIRCTWIFSOASF 391  
DB 298 EVKEMKKVTIISVHEKGFIEIKPTFSOLE--AVNLHEVYKHEVVEVRAVPPRISMLKNNLTL 356  
QY 392 PCEQRGLEDDYGISKFCO-----HKNK-----PGEYIFYAENDDAOQTKMFTLN 435  
DB 357 -----IENLTITTDVEKIQEIRYRSKLIRAKEEDSGHYTIVAEADAVKSYFELL 410  
QY 436 IKKPKQVL-----ANASASQSCSSDGYPLPSWTWKKCSKSPNCTEIPBPGVWNKKA 488  
DB 411 TQVPSIILDLVDHGGSTGGTAVRGTAEGLPLDIEWMICKD--IKKCNNETS---WTILA 466  
QY 489 NRKVGQWVSS-----STLNMSEAGKLLVKCCAYNSMGTSCTEIFLNSPG 534  
DB 467 NN-----VSNITTEIHSRDRSTVEGRVTFKAVEETIAVRCLAKNLLGAENRELKLV-- 518  
QY 535 PPFQIDNISFYATIGCLCLPFIWLVILVILCHYKQFRYESQLQMIQVTPGLDNEYFYVD 594  
DB 519 --PTLRSELTVAAAVLVLLVILVILVIVKQKPRYEIRWRVIESISPDGHEIYIYD 576  
QY 595 FRDYEDLKWEPPRENLEFGKVLGSGAFGRVNMATAYGISTGVSIQVAVKMLKEKADSC 654  
DB 577 PMQLPYDSRWKFPDRGLVLRVLSGAGKVGVEGTAYGLSRQPVKVMKVPKMTARSS 636  
QY 655 EKEALMSELKMTWTHLGHNDINVLGACTLGGPVYLIFFEYCCYDGLNLYLRSKRKEF--- 711  
DB 637 EKOALMSELKMTWTHLGHNLINVLGACTKSGPIIITEYCFYDGLVNVYLNHKNRDSFLSH 696  
QY 712 -----HRTWTEIFEKHNFSYPTFOAHNSNMP--GSREVQLHPP 749  
DB 697 HPEKPKKELDIFGLNPADSTRSYVLSFENGDYMDMKQADTTQYVPMLEKREVSYSKD 756  
QY 750 LDQSGFNGNSIHSEDEIYENOKRLAEEDLNVLTFEDLLCFAYQVAKGMFELEFS 809  
DB 757 IORSLYDRPASYKKSMLESDSEKLNLLSDNSEGLTLL--DLSFTYQVARGMEFLASKN 813  
QY 810 CVHRDLAARNVLVTHGKVVVKICDFGLARDILSDSYVVRGNARLPVKWMAPELFEIYT 869  
DB 814 CVHRDLAARNVLLAOGKIVKICDFGLARDIMHDSNVYSGSTFLPMKWMAPESIFDNLTY 873  
QY 870 IKSDVMSYGILLWEIFSLGVNYPGIPVDANFYKLIQSGFKMEQPFYATEGIYFVMQSCW 929  
DB 874 TISDVMWSYGILLWEIFSLGTPYPCGMVDSFTYFNKIKSGYRMAKPDHATSEVYEIMVKCW 933  
QY 930 AFDSRKRPSFPNLTSL 946  
DB 934 NSEPEKRPSFYHLSEIV 950

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:27:39 ; Search time 26.4867 Seconds  
(without alignments)  
3600.506 Million cell updates/sec

Title: US-09-919-408-2  
Perfect score: 5264  
Sequence: 1 MRALAQRSDRRLLLVLSV.....RGGLRAQPQRQVKIHRERS 992

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5264	100.0	992	A39931	protein-tyrosine k
2	5102	96.9	1000	S18827	Flt3 protein - mou
3	4408.5	83.7	993	A36873	protein-tyrosine k
4	1258.5	23.9	980	TVCTMD	macrophage colony-
5	1247	23.7	941	TVNMVD	protein-tyrosine k
6	1232.5	23.4	978	A49814	protein-tyrosine k
7	1223	23.2	975	TVMSKT	protein-tyrosine k
8	1221.5	23.2	972	TVHUMD	macrophage colony-
9	1220	23.2	954	I51703	c-kit-related kina
10	1215	23.1	977	I45877	protein-tyrosine k
11	1210	23.0	978	S16385	macrophage colony-
12	1198.5	22.8	976	TVMSMD	macrophage colony-
13	1196	22.7	976	TVHUKT	protein-tyrosine k
14	1188.5	22.6	975	T30816	macrophage colony-
15	1181	22.4	960	JN0677	protein-tyrosine k
16	1157.5	22.0	1088	PFRTGA	platelet-derived g
17	1152.5	21.9	1089	PFHUGA	platelet-derived g
18	1132	21.5	1087	I51552	platelet-derived g
19	1123.5	21.3	1089	S33727	platelet-derived g
20	1098	20.9	1098	PFMSRB	platelet-derived g
21	1078	20.5	1106	PFHUGB	platelet-derived g
22	1060	20.1	1048	T30815	platelet-derived g
23	998.5	19.0	1338	S09982	protein-tyrosine k
24	984.5	18.7	1336	I60598	receptor tyrosine
25	983.5	18.7	1333	I78875	Fit-1 tyrosine kin
26	980	18.6	1379	JC4954	vascular endotheli
27	970	18.4	1330	S49010	gag-kit polypeptid
28	967.5	18.4	790	F0MVH2	vascular endotheli
29	965	18.3	1348	S51656	vascular endotheli

ALIGNMENTS

RESULT 1

A39931  
protein-tyrosine kinase (EC 2.7.1.112) flk-2, hematopoietic - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jan-1992 #sequence\_revision 17-Jan-1992 #text\_change 04-Feb-2000  
C:Accession: A39931  
R:Matthews, W.; Jordan, C.T.; Wiegand, G.W.; Pardoll, D.; Lemischka, I.R.  
Cell 65, 1143-1152, 1991  
A:Title: A receptor tyrosine kinase specific to hematopoietic stem and progenitor cel  
A:Reference number: A39931; MUID:91292518; PMID:1648448  
A:Accession: A39931  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-992 <NAT>  
A:Cross-references: GB:M64689; NID:g193327; PIDN:AAA37634.1; PID:g193328  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom  
C:Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein  
F:609-953/Domain: protein kinase homology <KIN>  
F:617-625/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 5264; DB 2; Length 992;  
Best Local Similarity 100.0%; Pred. No. 4.7e-262;  
Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

30	954.5	18.1	1298	2	A48999	protein-tyrosine k
31	954.5	18.1	1356	2	JC1402	protein-tyrosine k
32	952	18.1	1363	2	I58375	protein-tyrosine k
33	926.5	17.6	1367	2	A41228	protein-tyrosine k
34	822	15.6	160	2	A39061	protein-tyrosine k
35	815	15.5	813	1	A49123	fibroblast growth
36	810.5	15.4	823	2	B35963	protein-tyrosine k
37	798	15.2	797	2	S38579	fibroblast growth
38	791	15.0	821	1	TVHUF2	fibroblast growth
39	791	15.0	824	2	S24108	protein-tyrosine k
40	789	15.0	806	2	A35963	protein-tyrosine k
41	788.5	15.0	769	2	S16236	fibroblast growth
42	788.5	15.0	822	2	A45081	fibroblast growth
43	780.5	14.8	822	2	A41794	keratinocyte growt
44	779.5	14.8	457	2	S44269	platelet-derived g
45	777	14.8	821	1	TVMSBK	fibroblast growth

```
QY 421 AENDDAQFTKMFNLIRKPKQVLANASASQASCSGYPPLPSWTWKKCDKSPNCTEIP 480
|||||
DB 421 AENDDAQFTKMFNLIRKPKQVLANASASQASCSGYPPLPSWTWKKCDKSPNCTEIP 480
QY 481 EGVWKKANKRVGQWSSSTLNNSAGKGLLVKCCAYNSMGTSCTIFLNSPGPFPIQ 540
|||||
DB 481 EGVWKKANKRVGQWSSSTLNNSAGKGLLVKCCAYNSMGTSCTIFLNSPGPFPIQ 540
QY 541 DNLSFYATIGLCIPFIVLVILVILCHIKKQFYESQLOMIQVTPDNEFYVDFRDY 600
|||||
DB 541 DNLSFYATIGLCIPFIVLVILVILCHIKKQFYESQLOMIQVTPDNEFYVDFRDY 600
QY 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISKTVSIQVAVKMLKEKADCEALM 660
|||||
DB 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISKTVSIQVAVKMLKEKADCEALM 660
QY 661 SELKMMTHLGHHDNIVNLLGACTLSGPVYLIFECYCGDLLNLYLRKREKPHRTWEIFK 720
|||||
DB 661 SELKMMTHLGHHDNIVNLLGACTLSGPVYLIFECYCGDLLNLYLRKREKPHRTWEIFK 720
QY 721 EHNFSYPTFOAHSNSMPSGSRVQLHPDQLSGFNGNSIHSEDETEYENQKRLAEDEE 780
|||||
DB 721 EHNFSYPTFOAHSNSMPSGSRVQLHPDQLSGFNGNSIHSEDETEYENQKRLAEDEE 780
QY 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
|||||
DB 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
QY 841 SDSYVYVGRNARLPVKWMAPELFEIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900
|||||
DB 841 SDSYVYVGRNARLPVKWMAPELFEIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900
QY 901 FYKLIOGFKMEOPFYATEGIYFVMSQWAFDSKRPSPFNLTSLGCOLAEAEACIRT 960
|||||
DB 901 FYKLIOGFKMEOPFYATEGIYFVMSQWAFDSKRPSPFNLTSLGCOLAEAEACIRT 960
QY 961 SIHLPKQAAPQQRGLRAQSPQKVIHRS 992
|||||
DB 961 SIHLPKQAAPQQRGLRAQSPQKVIHRS 992

RESULT 2
SIHLPKQAAPQQRGLRAQSPQKVIHRS 992
C:Species: Mus musculus (house mouse)
C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 24-Sep-1999
C:Accession: S18827
R:Rosnet, O.; Marchetto, S.; deLapeyriere, O.; Birnbaum, D.
A:Title: Murine Flt3, a gene encoding a novel tyrosine kinase receptor of the PDGFR/CSF1
Oncogene 6, 1641-1650, 1991
A:Reference number: S18827; MUID:92019834; PMID:1656368
A:Accession: S18827
A:Molecule type: mRNA
A:Residues: 1-1000 <ROS>
A:Cross-references: EMBL:X59398; NID:g50978; PID:CAA42041.1; PID:g50979
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: ATP
F:609-953/Domain: protein kinase homology <KIN>
F:617-625/Region: protein kinase ATP-binding motif

Query Match 96.9%; Score 5102; DB 2; Length 1000;
Best Local Similarity 96.7%; Pred. No. 9.2e-254;
Matches 968; Conservative 6; Mismatches 17; Indels 10; Gaps 2;
QY 1 MRALQSRDRLLLVLSVMILETVTNQDLPVIKCVLISHENNGSSAGKPSRYMVRGS 60
|||||
DB 1 MRALQSRDRLLLVLSVMILETVTNQDLPVIKCVLISHENNGSSAGKPSRYMVRGS 60
QY 61 PEDLOCTPRQSEGTVEATVEAAGSITLQVQLATPGDLSCLVWFKHSSLCQPHFD 120
|||||
DB 61 PEDLOCTPRQSEGTVEATVEAAGSITLQVQLATPGDLSCLVWFKHSSLCQPHFD 120
QY 121 LQNRGIVSMAILNVTQAGEYLLHIOSEANVTLVETVNVDRDTQLVLRPYFKMENQ 180
|||||
```

```
DB 121 LQNRGIVSMAILNVTQAGEYLLHIOSEANVTLVETVNVDRDTQLVLRPYFKMENQ 180
|||||
QY 181 DALLCISEGPEPTVEWVLCSSHRESCKEGPAVVRKEEVLHFGTDIRCCARNALGR 240
|||||
DB 181 DALLCISEGPEPTVEWVLCSSHRESCKEGPAVVRKEEVLHFGTDIRCCARNALGR 240
QY 241 ECTKLFITDLNQAPQSTLPQLFKVGEPLMIRKAIHVNHGFGLTWELEDALEEGSYFE 300
|||||
DB 241 ECTKLFITDLNQAPQSTLPQLFKVGEPLMIRKAIHVNHGFGLTWELEDALEEGSYFE 300
QY 301 MSTYSTNRTMIRILLAPVSSVGRNDTGYTCSSSKHPSQSALVTLLEKGFNATSSQBEY 360
|||||
DB 301 MSTYSTNRTMIRILLAPVSSVGRNDTGYTCSSSKHPSQSALVTLLEKGFNATSSQBEY 360
QY 361 EIDPYEKFCSVRFKAYPRICRTWIFSOASFPCEORGLDGYISKFCDHKNKPGYEIFY 420
|||||
DB 361 EIDPYEKFCSVRFKAYPRICRTWIFSOASFPCEORGLDGYISKFCDHKNKPGYEIFY 420
QY 421 AENDDAQFTKMFNLIRKPKQVLANASASQASCSGYPPLPSWTWKKCDKSPNCTEIP 480
|||||
DB 421 AENDDAQFTKMFNLIRKPKQVLANASASQASCSGYPPLPSWTWKKCDKSPNCTEIP 480
QY 481 EGVWKKANKRVGQWSSSTLNNSAGKGLLVKCCAYNSMGTSCTIFLNSPGPFPIQ 540
|||||
DB 481 EGVWKKANKRVGQWSSSTLNNSAGKGLLVKCCAYNSMGTSCTIFLNSPGPFPIQ 540
QY 541 DNLSFYATIGLCIPFIVLVILVILCHIKKQFYESQLOMIQVTPDNEFYVDFRDY 600
|||||
DB 541 DNLSFYATIGLCIPFIVLVILVILCHIKKQFYESQLOMIQVTPDNEFYVDFRDY 600
QY 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISKTVSIQVAVKMLKEKADCEALM 660
|||||
DB 601 DLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISKTVSIQVAVKMLKEKADCEALM 660
QY 661 SELKMMTHLGHHDNIVNLLGACTLSGPVYLIFECYCGDLLNLYLRKREKPHRTWEIFK 720
|||||
DB 661 SELKMMTHLGHHDNIVNLLGACTLSGPVYLIFECYCGDLLNLYLRKREKPHRTWEIFK 720
QY 721 EHNFSYPTFOAHSNSMPSGSRVQLHPDQLSGFNGNSIHSEDETEYENQKRLAEDEE 780
|||||
DB 721 EHNFSYPTFOAHSNSMPSGSRVQLHPDQLSGFNGNSIHSEDETEYENQKRLAEDEE 780
QY 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
|||||
DB 781 EDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKICDFGLARDIL 840
QY 841 SDSYVYVGRNARLPVKWMAPELFEIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900
|||||
DB 841 SDSYVYVGRNARLPVKWMAPELFEIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDAN 900
QY 901 FYKLIOGFKMEOPFYATEGIYFVMSQWAFDSKRPSPFNLTSLGCOLAEAEACIRT 960
|||||
DB 901 FYKLIOGFKMEOPFYATEGIYFVMSQWAFDSKRPSPFNLTSLGCOLAEAEACIRT 960
QY 961 SIHLPKQAAPQQRGLRAQSPQKVIHRS 992
|||||
DB 961 MGVNPEHPSIYQNRRLPSREAGS-EPPSPQAQKVIHRS 1000
|||||

RESULT 3
A36873
protein-tyrosine kinase (EC 2.7.1.112) STK-1 precursor - human
N:Alternate names: stem cell tyrosine kinase 1
C:Species: Homo sapiens (man)
C:Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 04-Feb-2000
C:Accession: A36873
R:Small, D.; Levenstein, M.; Kim, E.; Carow, C.; Amin, S.; Rockwell, P.; Witte, L.; B
Proc. Natl. Acad. Sci. U.S.A. 91, 459-463, 1994
A:Title: STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in CD34(+)
A:Reference number: A36873; MUID:94119906; PMID:7507245
A:Accession: A36873
A>Status: preliminary
```

A:Molecule type: mRNA  
A:Residues: 1-993 <SMA>  
A:Cross-references: GB:U02667  
A:Note: in the authors translation, an additional residue Ala is shown after 420-Ala and  
C:Genetics:  
A:Map position: 13q12  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: ATP; glycoprotein; phosphotransferase; transmembrane protein; tyrosine-speci  
F:608-950/Domain: protein kinase homology <KIN>  
F:616-624/Region: protein kinase ATP-binding motif

Query Match 83.7%; Score 4408.5; DB 2; Length 993;  
Best Local Similarity 84.0%; Pred. No. 2.8e-218;  
Matches 836; Conservative 57; Mismatches 89; Indels 13; Gaps 6;

QY 1 MRALAQRSDRLLLVLLVLSMILETVTQDLPVVKCVLISHENNGSSAGKSSYRWGRGS 60  
DB 1 MPALA-RDAGTVPLLVLFSAFIFGTITQDLPVVKCVLISHENNGSSAGKSSYRWGRGS 59  
QY 61 PEDLQCTPRQSEGVYEAATVEVAESGSILOVOLATPGDLSCVLWPKHSSGLGQPHFD 120  
DB 60 PEDLGCALRPOSGTVYEAAYVDVDSITLOVLVDAPGNISCLWPKHSSGLGQPHFD 119  
QY 121 LQNRGIYSMAILNVTQAGYLLHIOQSERANYTVLFTVNVVDTOLYVLRPFRKMNQ 180  
DB 120 LQNRGVVSMVLLKMTQTOAGYLLFIQSEATNYTLFTVSVIRNTLLYLRPFRKMNQ 179  
QY 181 DALLCISEGVPEPTVWVLCSSHRESCKEKPAPVYKKEKVLHELFGTDIRCCARNALGR 240  
DB 180 DALYCISESVPETVWVLCSSHRESCKEKPAPVYKKEKVLHELFGTDIRCCARNALGR 239  
QY 241 ECTKLFTDLNOAQSTLPQLFLKVGPELWIRCKAIHNVHGFGLTWELEDKALEGSYFE 300  
DB 240 ECTRLFTIDLNOTFOTLLPQLFLKVGPELWIRCKAHVHNGFGLTWELEDKALEGSYFE 299  
QY 301 MSTYSTNRTMIRILLAFVSSYGRNDTGYTCTSSSKHPQSQALVTILEKGFINATSSOBEY 360  
DB 300 MSTYSTNRTMIRILFAFVSSVARNDTGYTCTSSSKHPQSQALVTIVKGFINATSSSEY 359  
QY 361 EIDYEFKFCFVRKAPRIRCTWIFQASPCBORGLEDGYSTKFCDHKNKGEYIFY 420  
DB 360 EIDQYEFKFCFVRKAPRIRCTWIFSRKSPCEQKGLDNGYSISKFCNHHKQGEYIFY 419  
QY 421 -AENDDAQFTKMTLNTRKKPOVLNANASASQSSDGYPLPSWTWKKCDKSPNCTEEI 479  
DB 420 -AENDDAQFTKMTLNTRKKPOVLNANASASQSSDGYPL-SWTWKKCDKSPNCTEEI 478  
QY 480 PEGYNNKANKRVFGQWSSSTLNNKSPAGKGLLVKCCAYNSMGTSCEITFNPSGPPFFI 539  
DB 479 TEGYNNKANKRVFGQWSSSTLNNKSPAGKGLLVKCCAYNSMGTSCEITFNPSGPPFFI 538  
QY 540 QDNISFYATIGLCPFIWLVILVILCHYKKQFRYESQLOMIOVTGPDNNEYFYVDFRDE 599  
DB 539 QDNISFYATIGVCLLFIWLVILVILCHYKKQFRYESQLOMIOVTGPDNNEYFYVDFRDE 598  
QY 600 YDLKWEPPRENLEFGKVLGSAFGRVNNATAYGSKTGVSIOVAVKMLKERADSEKAL 659  
DB 599 YDLKWEPPRENLEFGKVLGSAFGRVNNATAYGSKTGVSIOVAVKMLKERADSEKAL 658  
QY 660 MSELKMTLGHHDHNIIVNLGACFTLSGPVYLIFCYCCVGDLLNLYRSKREKFRHTWTIF 719  
DB 659 MSELKMTQLSGHENIIVNLGACFTLSGPVYLIFCYCCVGDLLNLYRSKREKFRHTWTIF 718  
QY 720 KEHNFSSYPTFOAHSSNMPGSRVQLHPDQLDQSGFNNGSIHSEDETEYENQKRLABEE 779  
DB 719 KEHNFSSYPTFOAHSSNMPGSRVQLHPDQSDQISGLHNGSFHSEDETEYENQKRL--EE 776  
QY 780 EEDLVNLTFFEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLVTHGVKVKICDFGLARDI 839  
DB 777 EEDLVNLTFFEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLVTHGVKVKICDFGLARDI 836  
QY 840 LPSDSSVVRGNARLPVKWMAPESEFEGYITIKSDVWSYGILLWEIFSLGVNYPGIPVDA 899  
DB 837 MSDSNVVRGNARLPVKWMAPESEFEGYITIKSDVWSYGILLWEIFSLGVNYPGIPVDA 896  
QY 900 NPYKLIQSGFKWEQPFYATGEGYIFVWQSWAFDSRRKPSFPNLTSLGCOLAEEAC-- 957  
DB 897 NPYKLIQSGFKWEQPFYATGEGYIFVWQSWAFDSRRKPSFPNLTSLGCOLAEEAMYQ 956  
QY 958 -----IRTSIHLKQAAPOORG-GLRAQSPQORVK 986  
DB 957 NVDRGVSECPHYQNRPRFSREMDLGLLSPQAQVE 991

RESULT 4  
TVCTWD

macrophage colony-stimulating factor 1 receptor precursor - cat  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) csflr/fms  
C:Species: Felis silvestris catus (domestic cat)  
C>Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 13-Jun-1997  
C:Accession: A31636  
R:Woolford, J.; McAuliffe, A.; Rohrschneider, L.R.  
Cell 55, 965-977, 1988  
A:Title: Activation of the feline c-fms proto-oncogene: multiple alterations are requ  
A:Reference number: A31636; MUID:89077553; PMID:2849512  
A:Accession: A31636  
A:Molecule type: mRNA  
A:Residues: 1-980 <WOO>  
A:Cross-references: EMBL:X03663  
C:Genetics:  
A:Gene: fms  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo  
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming prote  
fic protein kinase  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-980/Product: macrophage colony-stimulating factor 1 receptor #status predicted <  
F:24-509/Domain: extracellular #status predicted <EXT>  
F:35-86/Domain: immunoglobulin homology <IMM1>  
F:120-179/Domain: immunoglobulin homology <IMM2>  
F:217-280/Domain: immunoglobulin homology <IMM3>  
F:316-381/Domain: immunoglobulin homology <IMM4>  
F:410-484/Domain: immunoglobulin homology <IMM5>  
F:510-534/Domain: transmembrane #status predicted <TM>  
F:535-980/Domain: intracellular #status predicted <INT>  
F:577-915/Domain: protein kinase homology <KIN>  
F:585-593/Region: protein kinase ATP-binding motif  
F:42-84,127-177,224-278,417-482/Disulfide bonds: #status predicted  
F:45,73,94,153,275,303,335,410,477,490/Binding site: carbohydrate (Asn) (covalent) #s  
F:613,630,776/Active site: Lys, Glu, Asp #status predicted  
F:781,794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.9%; Score 1258.5; DB 1; Length 980;  
Best Local Similarity 32.7%; Pred. No. 4.3e-57;  
Matches 328; Conservative 168; Mismatches 353; Indels 153; Gaps 32;

QY 75 TVYEAAATVEVAESGSILOVOLATPGDLSCVLWPKHSSGLGQPHFDLQNRG 125  
DB 13 TAWHAQGVVPVQPSGPELVPEPTVTLRCVNGSVEMDGPISPHNLDLDPSSI---- 68  
QY 126 IVSMAILNVTQAGEYLLHI-----QSERANYTVLFTVNVVDTOLYVL--RRPY----- 173  
DB 69 ---LFTNNATQNTCTY--HCTEPGNPGGNATI-----HLVKKDPAWPVKVLAQE 114  
QY 174 FRKMENQDALL-CISEGVPEPTVWVLCSSHRESCKEKPAPVYKKEKVLHELFGTDI-- 230  
DB 115 VTVLEGQDALLPCL---LTDPALE-----AGVSLVRVRGRPVLRQTNYSFSPWHGFTIHK 166  
QY 231 -----RCCARNALGRECTKL-----FTDILNQAPOSTL-PQLFLKV-GEPLWIRCK 274  
DB 167 AKFIENHVYQCSAR-VDGRTVTSMGILWKVQKDISGAPATLTLEPAELVRIQGEAAQIVCS 225  
QY 275 AIHVNHGFLWELEDKALEGSYFEMSTYSNRTMIRILLAFVSSVGRNDTGYTCSSS 334  
DB 226 ASNDIVNFDVSLRHGDTKL---TISQSDFDHNRVQ-KVLTNLNHDVFSQDAGNYSCTAT 281  
QY 335 K---HPSQSALVTILEKGFINATSSQE-EYETDPYEKFCFSVRFRKAYPRIC-TWI---- 385









A:Accession: I59083  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 874-972 <RE3>  
A:Cross-references: GB:M14193; NID:g182521; PIDN:AAA35834.1; PID:g182522  
R:Nienhuis, A.W.; Bunn, H.F.; Turner, P.H.; Gopal, T.V.; Nash, W.G.; O'Brien, S.  
Cell 42, 421-428, 1985  
A>Title: Expression of the human c-fms proto-oncogene in hematopoietic cells and its del  
A:Reference number: 152772; MUID:85282599; PMID:4028159  
A:Accession: I52772  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 244-295 <RE4>  
A:Cross-references: GB:M11067; NID:g182674; PIDN:AAA35848.1; PID:g442423  
C:Genetics:  
A:Gene: GDB:CSF1R; FMS  
A:Cross-references: GDB:I20600; OMIM:164770  
A:Map position: 5q33.2-5q33.3  
A:Introns: 17/1; 103/1; 198/1; 243/3; 297/1; 361/2; 400/1; 440/2; 504/1; 542/3; 585/1; 6  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;  
fic protein kinase  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-972/Product: macrophage colony-stimulating factor 1 receptor #status predicted <MAT  
F:24-512/Domain: extracellular #status predicted <EXT>  
F:35-86/Domain: immunoglobulin homology <IMM1>  
F:120-179/Domain: immunoglobulin homology <IMM2>  
F:217-280/Domain: immunoglobulin homology <IMM3>  
F:316-383/Domain: immunoglobulin homology <IMM4>  
F:412-487/Domain: immunoglobulin homology <IMM5>  
F:513-537/Domain: transmembrane #status predicted <TMM>  
F:538-972/Domain: intracellular #status predicted <INT>  
F:580-917/Domain: protein kinase ATP-binding motif  
F:588-596/Region: protein kinase ATP-binding motif  
F:42-84, 127-177, 224-278, 419-485/Disulfide bonds: #status predicted  
F:45, 73, 153, 240, 275, 302, 335, 412, 428, 480/Binding site: carbohydrate (Asn) (covalent)  
F:616, 633, 778/Active site: Lys, Glu, Asp #status predicted  
F:783, 796/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.2%; Score 1221.5; DB 1; Length 972;  
Best Local Similarity 32.5%; Pred. No. 3.3e-55;  
Matches 330; Conservative 152; Mismatches 331; Indels 201; Gaps 33;

QY 76 VYEATVEAESGSIITLQVQLATPGDLSCLVFKHSSLCGCPHFQDQNGRIVSMAL--- 132  
Db 22 VIEPSVPELVKPGAT--VTLRCVGVSGVENDGPPS-----PHWTLYSDG--SSILSTN 72

QY 133 NVTTQAGEYLLHQSERANTVLTFTVNVVDTQLYVLRPY-----FRKMENQDALL-C 185  
Db 73 NATFQNTGYRCTEFGDPLGGSAAIHLVYKDP-----ARPNNVLAQVGVVPEQDQALLPC 127

QY 186 ISEGVPEPTVEWVLCSSHRESCKEEGPAVVRKEE-----VLHE---LF 226  
Db 128 L---LTDPLV-----EAGVSLVRVGRPLMRHTNYSFSPWHGFTIHRAKFIQ 171

QY 227 GTDIRCCARNALGRECTKLTIDLNOAQPSLTLP-----QLFLKVGPELWTRCAI 276  
Db 172 SDQYQCSALMG-GR---KVMISITELKQKVPGPALTLPVPAELVRIRGEAAQIVCSAS 227

QY 277 HYNHGFGLTWLEDKALEEGSYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYITCSSS-- 334  
Db 228 SDVNFDFVLQHNNTKL--AIPOQSDPHNNRYQ-KVLTNLDQVDQFQHAGNYSVCAVSNV 283

QY 335 --KHPSQSALVILEKGFINTATSQE-EYEDIDPEKFCFSEVRFKAYPRIR-CTWIFSOAS 390  
Db 284 QGKH-STSMFFRVVSEAYISLSEONLLQEVTVGEGNLKVMVEAYPGILOGFNWY---- 338

QY 391 FPCGORGLEDGYSTSKFCDHKNKP-----GEYIFYAEN 423  
Db 339 -----LGPSSDHOPEKLANATTKDYRHTFTLSLPKLPSEAGRYSFLARN 385

QY 424 DDAOFTKMTLNIKKPQV-----LANASASQACSSDGYPLPSPWTKKCSKSPNCTEE 478

Db 386 PGGWRALTEELTLRYPPEVSVIWTFFINGSGT--LLCAASGYPPQPNVTWLQCSGHTDRDEA 444  
QY 479 IPEGVWN---KKANKKVFQWVSSSTLNMSSEAGKLLVKKCAYNMSGTCETIFLSPG 534  
Db 445 QVLQWDDDPYEVLSQEPFHVKTVQSLTIVTETLEHNOTYECRAHNSVSGSWAFIPISAG 504  
QY 535 PPFIODNISFYATIGLC---LPFIVVILVILCHIKYKKQFRYESQLOMIOVGTGPLDNEYF 591  
Db 505 ANTHPPDEFLFTPVVAVACHSIMALLLLLLLLLYKKQPKYQVRWAIIE--SYEGNSYT 562  
QY 592 YVDFRDEYDLKWEFFPRENLEPKVGLSGSAFGRVMNATAYGISTGTVSIOVAVKMLKEKA 651  
Db 563 FIDPTOLPYNEKWEFFPRNNLQFGKTLGAGAFGVKVEATAFGLKEDAVLVAVKMLKSTA 622  
QY 652 DSCKEALMSELKMTHTLGHNDIVNLGACTLSGPVYLIFEYCCYGDLLNLYLSKREKF 711  
Db 623 HADEKALMSELKIMSHLGOHENIVNLGACTHGGPVLVITEYCCYGDLLNLYLSKREKF 680  
QY 712 HRTWTEIFKEHNFSSYPTFEQAHSSNMPGSRREVLPPLDQLSGFNGNSIHSEDE---- 766  
Db 681 -----AMLGF---SLSPGQDPEGVDYKNHLEKYYVRD 712

QY 767 -----IYENOKRLADE--EEDLNVLTFEDLLCFAYQVAKMEFLFK 808  
Db 713 SGFSQGVDTYVEMRDPVSTSSNDSFSEQDLDKEDGRPLELRLDLLHFSQVQAQMAFLASK 772

QY 809 SCVHRDLAARNVLVTHGVKVKICDFGLARDILSDSSVYVVRGNARLPVKWMAPELSLFGIY 868  
Db 773 NCIHRRDVAARNVLLTNGHVAKIGDGLARDIMNDSNIYKGNARLPVKWMAPELSIFDCVY 832

QY 869 TIKSDWVSYGILLWEIFSLGVNVPQGPVDANFYKLIQSIOGKMEQPFYATGTYFVWOSC 928  
Db 833 TVQSDWVSYGILLWEIFSLGLNYPGILVNSKEYKLKVDGYQMAQAFAPAFKNYISIMQAC 892

QY 929 WAFDRKRKPSFNLTSLFQCCLQAEAEACIRTSIHLPKQAAPQOGRGLRAQSPQ 982  
Db 893 WALEPHTHRTFQOICSFL--QEQAQEDRRERYDTNLPSSS---RSGSGSGSSSE 941

RESULT 9  
151703  
C:kit-related kinase 1 (Xkrk1) - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 18-Jun-1999  
C:Accession: I51703  
R:Baker, C.V.; Sharpe, C.R.; Torpey, N.P.; Heasman, J.; Wylie, C.C.  
Mech. Dev. 50, 217-228, 1995  
A>Title: A Xenopus c-kit-related receptor tyrosine kinase expressed in migrating stem  
A:Reference number: I51703; MUID:95344996; PMID:7619732  
A:Accession: I51703  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-954 <BAK>  
A:Cross-references: EMBL:248770; NID:g763033; PIDN:CAA88688.1; PID:g763034  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo  
C:Keywords: ATP  
F:575-915/Domain: protein kinase homology <KIN>  
F:583-591/Region: protein kinase ATP-binding motif

Query Match 23.2%; Score 1220; DB 2; Length 954;  
Best Local Similarity 32.0%; Pred. No. 3.9e-55;  
Matches 315; Conservative 160; Mismatches 321; Indels 188; Gaps 35;

QY 77 YEATVEAESGSIITLQVQLATPGDLSC-----LWVFKHSSLCGCPHFQDQNGRIVSMA 130  
Db 15 YTGDAVPKINDGEDRVTVNVGVKVSLECDRAHLVTLAFQKSLGMLKKPR-DLKSRLP----- 68

QY 131 ILNVTTET-----QAGEYLLHIOSEARANTVLTFTVNVVDTQLYVLRPYRPFKME 178  
Db 69 -LNNSETDOFFVIKADLRHIGRYICTNTQENTSV--SLFKVDPAPELDFIPFDVTE 125

QY 179 NODAL-LCISEGVPEPTVEWVLCSSHRESCKEEGPAVVRKEEVLHLEFGTDTRC----- 232

Db 126 GADTVGMCFP---TDPDMDBIAI-----EKC---DGSPLPE-----NFTFTTDIEAGITIK 169  
Qy 233 -----CARNALGR-ECTKLTIDLNOAPOSTLPQLFL-----KYGEPLWIRC 273  
Db 170 TVQLAFDSCYVCGNKGSGTGVKSSSTFISHVKVPKK-VPTVELSKSRQLVKGEPEVTC 228  
Qy 274 KAIHVNHGFGLTW-ELEOKALBEGSYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYTCS 332  
Db 229 AVLDFVSTVKAQMLDVKEGVTQANFRSSNVFSNLTAKS-----DGVPSRESRTFTTCQ 282  
Qy 333 SKKHPQS---SALVTILEKGFNATSSQP-EYEIDPYEKFCFSVRKAYPR-IRCTWIFS 387  
Db 283 AENAIQOVNATFTLVIDVGGYVNLVLENTTISVNAAGDLVLVKVYIDAYPDGQVWTF 342  
Qy 388 QASFPCEQRLGLEDYSISKFCDHKN-----KPEGEIFYAENDDAQTKMF 432  
Db 343 NETL-----LNTSDHYVATK--DEGNRRYVSELHLRLKTEKGVTYFTTNSDDASVSF 396  
Qy 433 TLNIRKKPQVLANASASQAS--CSSDGYPLPSWTWKKCDKSPNCTEETPEGVWKK--A 488  
Db 397 NIQVTRPEILLAERTSEGTLOCVATGFPVPAIQWYFCPSQORCTDYPPLSPVNEKFQ 456  
Qy 489 NRKVGQWSSSTLMSERAGKGLLVKCCAYNSMGTS-----CETIFLNSPQP 535  
Db 457 ENSSLGRIVVESTIDVNDLKNKGTVCQVASNEVESAYSFSAIKKELRTHLTFT----- 511  
Qy 536 PFIQDNTISFYATIGLCPLPFIIVLVILCHIKKYKOPRYESOLOMI-OVTPGLDNEYFYVD 594  
Db 512 -PLL---IGFIAAAGL---MCIAVAVLMYKYLQPKYEIQWKVVEEING---NNTVYID 560  
Qy 595 PRDYEYDLKWEPPRENLEFGKVLGSGAFGRVMNATAYGISKTSQVAVKMLKADSK 654  
Db 561 PTQLPYDNKWEPPRDLCEGKILGAGAFKVEATAYGLLKEDSRILTAVKMLKPSAHT 620  
Qy 655 EKEALMSELKMTLGHHDNIVNLGACTLSGPVYLIFCYCYGDLNLYLRSKREKFRHT 714  
Db 621 EREALMSELKVLSLGHKHNIVNLGACTVGGPTLVITEYCCYCGDLNLYLRSKRDSEF--- 677  
Qy 715 WTEIFEKHNFSYPTFOAHNSMSPGSRREVOLHPPLDQLSGNGNSIHSDEIEYENOK- 773  
Db 678 -----ICPKFEDNS-----EALYKNL-----LNRDMGCEGMSYIDMKP 713  
Qy 774 -----RLAEEEEEDLNVLTFEDLLCFAYQVAKGMEFLEFKSC 810  
Db 714 AVSYVVPKTKRRSGSGDQDVSYSIPEDDL-ALDTEDLNFSQVQAQGMFLASKNC 772  
Qy 811 VHRDLAARNVLYTHGKVKVICKDFGLARDILSDSSYVVRGNARLPVKWMAPESEFYI 870  
Db 773 IHRDLAARNILLTHGRITKICDFGLARDIRNDSNYVVKGNARLPVKWMAPESEIFHCYTF 832  
Qy 871 KSDVWSYGILLWEIFSLGVNYPGIPVDANFYKLIQSGFKMEOPFYATGFIYVMQSCWA 930  
Db 833 ESDVWSYGILLWEIFSLGSSYPRIPIVDVSKFYKMIKDGYYRMSPECAPLEMWEIMRSCWN 892  
Qy 931 FDSRRKRPSPNLTSLFLGCOLAAE 954  
Db 893 SDPLKRPTEKQIVQWVEQQLSDSK 916

## RESULT 10

I45877  
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - aurochs  
C:Species: Bos primigenius (aurochs)  
C:Date: 19-Dec-1997 #sequence\_revision 19-Dec-1997 #text\_change 20-Apr-2000  
C:Accession: I45877  
R:Kubota, T.; Hirono, H.; Sakurai, M.  
Gene 141, 305-306, 1994  
A:Title: Sequence of a bovine c-kit proto-oncogene cDNA.  
A:Reference number: I45877; MUID:94215924; PMID:7512939  
A:Accession: I45877  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-977 <KUB>

A:Cross-References: GB:D16680; NID:9516659; PIDN:BA04084.1; PID:g516660  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo  
C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase  
F:329-395/Domain: immunoglobulin homology <IM>  
F:588-932/Domain: protein kinase homology <KIN>

## Query Match 23.1%; Score 1215; DB 2; Length 977;

Best Local Similarity 32.8%; Pred. No. 7.2e-55;  
Matches 304; Conservative 141; Mismatches 298; Indels 184; Gaps 28;

Qy 134 VTE----TOAGEYLLHIQSERANYTLVFTVNVDRTO-LYVLRPRPYFKMENODALLCI-- 186  
Db 83 ITEKAEATNGTVC---TNKGGLSSIIYVVRDPEKFLDILDPYCKEENDTLVRCLPT 139  
Qy 187 -----SEGVPPEP-TVEMWLCSSHRESCKECPAVVRKEEKLHFLGFTDIRC-- 232  
Db 140 DEVTNYSLTGCEGKPLPKDLTFV-----ADPKAGITIRNVKREYHRL-----CLH 185  
Qy 233 CARNALGRE-CTKLTIDLNOAPOS-----TLPQLFKYGEPLWIRCKAIHVNHGELT 285  
Db 186 CSANORGKSMLSKKFTLKVRAAIKAVPVSVSKTSYLLREGEFAVTCILIKDVSSVDSM 245  
Qy 286 WELEDKALBEGSYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYTCSKSKHPSOSALVTI 345  
Db 246 WIKENSQOTKAQTKKNSWHQGDPSYLRQERLTISSARVNDGVMFCYANNTFGSANVTIT 305  
Qy 346 LE---KGFINA-TSSOEVEYIDPYEKFCFSVRKAYPR-IRCTWIFSQAAPPCQORGLD 400  
Db 306 LEVVDAGFINIFPMNTTVFVNDGENVDLVVEYEAYPEKVRHQWIYMNRT----- 355  
Qy 401 GYSISKFCDHK-----NKPGEYIFYAENDDAQTKMFTLNIRKK 439  
Db 356 --STDKWDYDYPKSENEINIRYVNLHLRLKGTGGTYTFHVSNSDVNSSTVFNVYNTK 413  
Qy 440 PQVLANASA--SOACSSDGYPLPSWTWKKCDKSPNCTEETI-PEGVWKNKARKVFGQW 496  
Db 414 PEILTHDLRVNGLMLQCAAGFPEPTIDWYFCPCQTEQRCSPVGPVDVQIQNSVSPSEGL 473  
Qy 497 VSSTLNMSBAGLLVKKCCAYNSMGTS-----CETIFLNSPQPFPFI 539  
Db 474 VVYSTIDSTFKINGTVECRAYNDVGKSSASFNAFKGNSKEQIHAHTLFT-----PLL 527  
Qy 540 QDNISEYATIGLCPLPFIIVLVILCHIKKYKOFRYESOLOMI-OVTPGLDNEYFYVDRDY 598  
Db 528 ---IGFVIAAGLCIFWML---TYKYLQKPMYEVQWKVVEEING---NNYVYIDPTQL 577  
Qy 599 EYDLKWEFPRENLEFGKVLGSGAFGRVMNATAYGISKTSQVAVKMLKADSKCEKEA 658  
Db 578 PYDHKWEFPKNRISFGKTLGAGAFKVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREA 637  
Qy 659 LMSLKMTHLGHHDNIVNLGACTLSGPVYLIFCYCYGDLNLYLRSKREKFRHTWTEI 718  
Db 638 LMSLKVSLVSLGHNMNIVNLGACTIGGPTLVITEYCCYCGDLNLYLRSKRDSEF----- 690  
Qy 719 FKEHNFSYPTFOAHNSMSPGSRREVOLHPPLDQLSGNGNSIHSDEI-----EYENQ 772  
Db 691 -----ICSKQEDHAEVALYK-----NLLHSKSSCNDSTNEYMDM 725  
Qy 773 K-----RLAEEEEEDLN-----VLTFEDLLCFAYQVAKGMEFLE 806  
Db 726 KPGSVYVVPKTKRRSARSIYERDVTPTAIMEDEDLALDELDSFSYQVAKGMFLA 785  
Qy 807 FKCVHRDLAARNVLYTHGKVKVICKDFGLARDILSDSSYVVRGNARLPVKWMAPESEIFEG 866  
Db 786 SKNCIHRDLAARNILLTHGRITKICDFGLARDIRNDSNYVVKGNARLPVKWMAPESEIFNC 845  
Qy 867 IYTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYKLIQSGFKMEOPFYATEGFIYFYHQ 926  
Db 846 VYTFESDVWSYGIFLWELFSLGSSYPGMPVDVSKFYKMIKEGFRMLSPHAPAEYDIK 905  
Qy 927 SCWAFDSRRKRPSPNLTSLFLGCOLAAE 953  
Db 906 TCWDADPLKRPTEKQIVQVLEKQISES 932

## RESULT 11

S16385  
macrophage colony-stimulating factor 1 receptor precursor - rat  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) CSF-1R  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 02-Jun-2000  
C:Accession: I60321; S16385  
R:Borycki, A.G.; Guillier, M.; Leibovitch, M.P.; Leibovitch, S.A.  
Growth Factors 6, 209-218, 1992  
A>Title: Molecular cloning of CSF-1 receptor from rat myoblasts. Sequence analysis and  
A:Reference number: I60321; MUID:93001225; PMID:1389227  
A:Accession: I60321  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-978 <RES>  
A:Cross-references: EMBL:X61479; NID:957543; PIDN:CAA43706.1; PID:957544  
A>Note: in Genbank entry RRGCSF1, release 113.0, the source is designated as Rattus rattus  
A>Note: submitted to the EMBL Data Library, August 1991  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
C:Keywords: ATP; autophosphorylation; glycoprotein; growth factor receptor; kinase-related  
protein; tyrosine-specific protein kinase  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-978/Product: macrophage colony-stimulating factor 1 receptor #status predicted <MAT>  
F:20-515/Domain: extracellular #status predicted <EXT>  
F:35-86/Domain: immunoglobulin homology <IMM1>  
F:120-179/Domain: immunoglobulin homology <IMM2>  
F:217-280/Domain: immunoglobulin homology <IMM3>  
F:316-381/Domain: immunoglobulin homology <IMM4>  
F:410-485/Domain: immunoglobulin homology <IMM5>  
F:516-535/Domain: transmembrane #status predicted <TM>  
F:536-978/Domain: intracellular #status predicted <INT>  
F:578-915/Domain: protein kinase homology <KIN>  
F:586-594/Region: protein kinase ATP-binding motif  
F:42-84, 127-177, 224-278, 417-483/Disulfide bonds: #status predicted  
F:45, 73, 302, 335, 389, 410, 449, 491/Binding site: carbohydrate (Asn) (covalent) #status  
F:614, 631, 776/Active site: Lys, Glu, Asp #status predicted  
F:781, 794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.0%; Score 1210; DB 2: Length 978;  
Best Local Similarity 33.5%; Pred. No. 1.3e-54;  
Matches 318; Conservative 147; Mismatches 329; Indels 154; Gaps 31;

QY 78 EAATVEVAESGSIITLOVQATPPGDLSCSLWFKHSSGCGPHDQNGRIVSMAILNVTF 137  
DB 36 ETVTLCRVCSNVEVD-----GPISPTWLDPSGSGS-----TLTTRNATPK 77  
QY 138 QAGEY-LLHIQSERANYTVLETVNVRDTQLVY-----LRRPYFRKMENODALL-CIS 187  
DB 78 NTGYRCTEEDPAGSTTI-----HLVYKDPAHWNLLAQEVTVVGEQEAIVPLCL- 128  
QY 188 EGVPEPTVEWVLCSSHRSCKEGPAVVRK-----EEKVLHELFGTDIRCC 233  
DB 129 --ITDPALK-----DSVSLMREGGROVLKTVYFSAWRGFTIRKAKVL-----DSNTVVC 177  
QY 234 ARNALGRECTKL-TTIDLN-----QAQSTL-POLELV-GEPLWTRCKAIHVNHGFLTW 286  
DB 178 KTMVNGRESTGTGLWLVNRVHPEPPQIKLEPKSLVIRGGAQAQVCSATNAEVEFNIL 237  
QY 287 ELEDKALE--EGSYFEMSTYSTNRTMIRILLAFVSSVGRNIDGYTC-SSSKHPQSALV 343  
DB 238 KRGTCKLEIPNSFDQNYKKVRL-----SLNAVDFDAGIYSCVASNDVGFRRTATM 291  
QY 344 T--ILEKGFINATSQE-EYEDIPYEKFCFSVRKAYPRIR-CTWIFSQAGPPCQORGLE 399  
DB 292 NFQVVEAYLNLTSQSLQVSGVDSILITVHADAYPSIQHNTYILGPFFE-DQKLE 350  
QY 400 -----DGYSISKPCDH--KNKPGEYIFYAENDDAQFTKMTLNRKKPQVLANASQ 450  
DB 351 FITORATRYRTFKLFLNRVKASEAGQYFLMAQNKAGWNLLTFELTLRYPPEVSVTWPVN 410  
QY 451 AS-----CSSDGYPLPSWTWKCKSDKSPNCTEIEPGVWN-----KKANRKYVQGVSSSTL 502

DB 411 GSDVLFCDVSGYPQPSVTWMECRGHTDRCDQAQALQVWNDTHPEVLQSKQFDPKVIQSQL 470  
QY 503 NNSEAGKGLLVKCCAYNSMGTSCTETIFLNSPGPPPIQDINSFYATIGLC---LPFIVVL 559  
DB 471 PIGTLKHNMTYFCKTHNSVGNSSQYFRAVSLGOSKOLPDESLETPVVVACMSVMSLLVL 530  
QY 560 IYLVICHKKYKOFRESOLOMIQ-VTGPLDNEYFYVDFRDYEDLKWEFFRENLEFGKVLG 618  
DB 531 LLLLLYKQKPKYQVWRKIIERYEG---NSYTFIDPTQLPYNEKWEFFPRNNLOFGKTLG 587  
QY 619 SGAFGRVMNATAYGISGTGVSIOYAVKMLKEKADSCKEALMSSELKMTLHGHHDINVLN 678  
DB 588 AGAFGKGVVEATAFGLGKEDAVLKAVYKMLKSTAHADEKEALMSSELKMTLHGHHDINVLN 647  
QY 679 LGACTLSGPVYLIFEYCCYGDLLNLYRSKRE-----KPHRTWTEI 718  
DB 648 LGACTHGGPVLVITEYCCYGDLLNLYRSKRE-----KPHRTWTEI 707  
QY 719 FKEHNFSSYPTFOAHSSNMPGSRVOLHPPDLQSLGFCNGNSIHSEDEIEYENQKRLAEE 778  
DB 708 RRDSGFS-----SQGVDTVYEMRPVSTSSSDSFFKQD-----L 740  
QY 779 EREDNLNLTFFEDLLCFAYOVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVVICDFGLARD 838  
DB 741 DKEPSRPLEMDLLHFFSSQVQAQMAFLAKNCIHRDVAARNVLLTSGHVAKIGDFGLARD 800  
QY 839 ILSDSYVVRGNARLVKVMAPESLFEIYTIKSDVMSYGILLWEITPSLVGNVYPGIPVD 898  
DB 801 IMNDSNVYVKGARLVKVMAPESILYCVTVQSDVMSYGILLWEITPSLVGNVYPGILVN 860  
QY 899 ANFYKLIOGSFKNQOPFYATEGIYFYVQSCWAFDSKRKPSFPNLTSL 946  
DB 861 NKPYKLVDGYQMAQPVFAKPNIIYSIMQSCWDLPTRRPTFQOICILL 908

## RESULT 12

TVMSMD  
macrophage colony-stimulating factor 1 receptor precursor - mouse  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) csflr/fms  
C:Species: Mus musculus (house mouse)  
C>Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 02-Jun-2000  
C:Accession: S01880  
R:Rothwell, V.M.; Rohrschneider, L.R.  
Oncogene Res. 1, 311-324, 1987  
A>Title: Murine c-fms cDNA: cloning, sequence analysis and retroviral expression.  
A:Reference number: S01880; MUID:88217329; PMID:2966922  
A:Accession: S01880  
A:Molecule type: mRNA  
A:Residues: 1-976 <ROT>  
A:Cross-references: EMBL:X06368  
C:Genetics:  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo  
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming prote  
fic protein kinase  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-976/Product: macrophage colony-stimulating factor 1 receptor #status predicted <EXT>  
F:20-515/Domain: extracellular #status predicted <EXT>  
F:35-86/Domain: immunoglobulin homology <IMM1>  
F:120-179/Domain: immunoglobulin homology <IMM2>  
F:217-280/Domain: immunoglobulin homology <IMM3>  
F:316-381/Domain: immunoglobulin homology <IMM4>  
F:410-485/Domain: immunoglobulin homology <IMM5>  
F:516-535/Domain: transmembrane #status predicted <TM>  
F:536-976/Domain: intracellular #status predicted <INT>  
F:578-914/Domain: protein kinase homology <KIN>  
F:586-594/Region: protein kinase ATP-binding motif  
F:42-84, 127-177, 224-278, 417-483/Disulfide bonds: #status predicted  
F:45, 73, 302, 335, 389, 410, 449, 478, 491/Binding site: carbohydrate (Asn) (covalent) #stat  
F:614, 631, 776/Active site: Lys, Glu, Asp #status predicted  
F:781, 794/Binding site: magnesium (Asn, Asp) #status predicted



A: Introns: 23/1; 113/1; 207/1; 252/3; 309/1; 372/2; 411/1; 449/2; 514/1; 549/3; 592/1; 6  
A: Note: defects in this gene may result in piebaldism  
C: Function:  
A: Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
C: Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
C: Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; kinase-related  
rotein; tyrosine-specific protein kinase  
F: 1-976/Product: protein-tyrosine kinase  
F: 1-509/514-976/Product: protein-tyrosine kinase kit precursor, long form #status predicted <MAT>  
F: 1-22/Domain: signal sequence #status predicted <SIG>  
F: 23-976/Product: protein-tyrosine kinase kit #status predicted <MAT>  
F: 23-520/Domain: extracellular #status predicted <EXT>  
F: 51-99/Domain: immunoglobulin homology <IMM1>  
F: 129-188/Domain: immunoglobulin homology <IMM2>  
F: 226-292/Domain: immunoglobulin homology <IMM3>  
F: 328-394/Domain: immunoglobulin homology <IMM4>  
F: 423-493/Domain: immunoglobulin homology <IMM5>  
F: 521-543/Domain: transmembrane #status predicted <TM>  
F: 544-976/Domain: intracellular #status predicted <INT>  
F: 597-931/Domain: protein kinase homology <KIN>  
F: 595-603/Region: protein kinase ATP-binding motif  
F: 58-97/136-186,233-290,428-491/Disulfide bonds: #status predicted  
F: 130,145,283,293,300,320,352,367,463,486/Binding site: carbohydrate (Asn) (covalent) #s  
F: 623,640,792/Active site: Lys, Glu, Asp #status predicted  
F: 797,810/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 22.7%; Score 1196; DB 1; Length 976;  
Best Local Similarity 31.4%; Pred. No. 6.8e-54;  
Matches 312; Conservative 172; Mismatches 361; Indels 148; Gaps 32;

QY 36 CVLISHENNGSACKPSYRMVGRSPEDLOCTPRQSEGTVEATVEAESGTTLOVQ 95  
DB 12 CVLLLLLRVQTGSQPS-----VSPGPPSPSIHPGKSDLIVRGD-EIRLL 57  
QY 96 LATPGDLSCLVAFKHSGLGCPHFDLQNRGIVSMALNVTTQAGEYLLHQSERANTV 155  
DB 58 CTDPGFVK--WTFE---ILDETENKQWETEK-----EATNTGKYCTTKRKHLSNIY 108  
QY 156 LFTVNRD-TOLYVLRPFYFKMENQDALLCI-----SEG-----VPE 192  
DB 109 VF---VRDPAKFLVDRSLYCKEDNDTLVRCPLDPEVNTSLKGCQCKPLPKDLRFPD 165  
QY 193 PTVEWLCSHRE-----SKKEGPVYVKEKVLHLEFGTDIRCCARNALGRECT 244  
DB 166 PKAGIMIKSVKRAYHRLCHLHCSVOEG-----KSLSEKFLKVRPAFK-----A 210  
QY 245 LFTIDLQAQSTLPQLFLKVGEPFLWTRCKAIHNVHNGFLTWLEEDK--ALEE----- 295  
DB 211 VPVSVSKA-----SYLLREGEFTVTCTIKDVSSSVYTWKRENSOTKLQEKYNWHH 264  
QY 296 GSYFEMSTYSTNRTMIRILLAFVSSVGRNDGYTCTSSSKHPSQALVTILE---KGFN 352  
DB 265 GDF-----NYEQATLT-----ISSARVDSGVFWCFWANTFTGSGANVTTLLEVVDKGFN 314  
QY 353 A-TSSQBEYEDIDPYKFCFSYRFRAYPR-IRCTVYFQASP--PCEQRGLDGYISKFC 408  
DB 315 IFPMINTTVFVNDGENVDLIVYEAFKPEHQOQYNNRTTDTKWDYTPKSENEINRYV 374  
QY 409 DHKN-----KPEYIFYAENDDAQTKMTLIRKQPQLVA--NASASQSSSDGY 459  
DB 375 SELHLRLKGTGGTYTFLVSNSDVNAIAFNVTNKPEITLDRLVNGMLQCVAAQFP 434  
QY 460 LPSMTWKCKSDKSPNCTEEI-PEGVWNNKANRKYFGQWSSVSTLNMSEAGKLLVKCCAY 518  
DB 435 EPTIDWFCGTEORCSASLPVDVQTLNNSGPPFGKLVVOSSIDSSAFKINGVECKAY 494  
QY 519 NSMGTSCETIFLNSPGPPFFIQDN-----ISFYATIGLCPLPIVVLVILIC 564  
DB 495 NDVGKT--SAYFN---FAFGNNKEQIHPTLTFTPLLIGIVVAGM-----MCIIVMLT 544  
QY 565 HKYKKOFYESQLQMI-OVTGPLDNEYFYVDREDYDLKWEFFRENLEFGKVLGSGAFG 623  
DB 545 KYVLQKPMYEQWKEVEING---NNVYIDPTQLPYDHWKWEFFRNRLSFGKTLGAGAFG 601

QY 624 RVMNATAYGISKTGVSIOVAVKMLKEKADSCKEALMSELKMMTHLGHHDNIVNLLGACT 683  
DB 602 KVVETATYGLIKSDAAMTVAVKMLKSAHLTEREALMSELKVLKSYLGNHNHIVNLLGACT 661  
QY 684 LSGPVYLFEYCCYCGDLLNLYLSRKREKHFRTWTEIFKEHNFSSYPTFOAHNSNSMPGSR 743  
DB 662 ICGPTLVITEYCCYCGDLLNLYLSRKREKHFRTWTEIFKEHNFSSYPTFOAHNSNSMPGSR 721  
QY 744 VOLHPLDOL---SGFNGNSIHSEDEIEYENQKRLAEEDENLVLTFFDLCLFAVQAK 800  
DB 722 MDMPKGVSVVPTKADKRVRIGSYIERDVTPAIMEDEDELALDL---EDLLSFSYQVAK 778  
QY 801 GMEFEFKSCVHRDLAARNVLVTHGKVKICDFGLARDILSDSSYVVRGNARLPVKMMAP 860  
DB 779 GMAFLASKNCIHRDLAARNLILTHGRTIKICDFGLARDIKNDNSYVVRGNARLPVKMMAP 838  
QY 861 ESLFEGIYTIKSDVMSYIGILLMEIFSLGVNYPGIPVDANFYKLIQSGFKMEOPFYATEG 920  
DB 839 ESIFNCVTFESDVMSYIGIFLWELFSLGSSPYGMPVDSKFYKMKIEGFRMLSPHAPAE 898  
QY 921 IYFVQSWAFDSRKRPPSPNLTSLGCOLAE 953  
DB 899 MYDIMKTCWDADPLKRPTFKQIVQLIEKQISES 931

RESULT 14  
T30816  
macrophage colony-stimulating factor receptor - Japanese pufferfish  
C: Species: Fugu rubripes (Japanese pufferfish)  
C: Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Jun-2000  
C: Accession: T30816  
R: How, G.F.; Venkatesh, B.; Brenner, S.  
Genome Res. 6, 1185-1191, 1996  
A: Title: Conserved linkage between the pufferfish (Fugu rubripes) and human genes for  
C: Reference number: 220882; MUID: 97129405; PMID: 8973913  
A: Accession: T30816  
A: Status: preliminary  
A: Molecule type: DNA  
A: Residues: 1-975 <HW>  
A: Cross-references: EMBL:U63926; NID: g1752706; PID: g1752708; PIDN: AAC50063.1  
C: Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo

Query Match 22.6%; Score 1188.5; DB 2; Length 975;  
Best Local Similarity 33.0%; Pred. No. 1.6e-53;  
Matches 298; Conservative 153; Mismatches 291; Indels 161; Gaps 25;

QY 164 TOLYVLRPFYFKMENQDALLCISEGVPEPTVWVLCSSHRESCKEEGPAAVVRKEEVLH 223  
DB 124 TSLRVVRK-----EGEDYLLCLLTDPATDLGLRMDNGTTVPPEMNYTVYRHRGILIR 177  
QY 224 EL---FGTDIRCCARNALGRECTKLTITDLNQAPOSTLPOLFLK-----VGEPLWIRC 273  
DB 178 SLOPSFNADYVCTAKVKGVEKTSKTSINVIQLRPP-PYVLEMDYEVYRIVGEELQIRC 236  
QY 274 KAIHNVHNGFGLTWLEDKA--LEEGSYFEMSTYSTNRTMIRILLAFVSSVGRNDGYTT 330  
DB 237 MTHNPFNNTVNTVNTTTSRVITIE---RVRSSENGENRLDIQSILT-ISAVDLADTGNIS 291  
QY 331 CSSSKHP---SOSALVTITILEKFI-----NAVSSQBEYEDIDPYKFCFSYRFRAYPR 379  
DB 292 CIGTNEAGVNSNTTLLVVEKPYIRLWQILPKLASOGLSVENEGEDLELGVWVAYPQ 351  
QY 380 IRCTWIFSOAFPCQORGLDGYISKFCDDHK-----TDRWHTPTSPSTSMQEHYHARLQKRMNAQ 384  
DB 352 I-----  
QY 413 KPEYIFYAENDDAQTKMTLIRKRP-QVLNANASASQSSSDGYPLPSWTWKCKSDK 471  
DB 385 EGOGYTFYAKNLANGLSISFHKYKQKPIAVVRWENITTLTCTSEFGYPAQIIWQSGI 444  
QY 472 SPNCT-----EIEPGVWNNKANRKYFGQWSSVSTLNMSEAGKLLVKCCAYNSMCT 523

```
Db 445 RPTCNCNGLPKQNHPOAL-TVEQREBEYGAIVESVFTVGLSNHRMTVECAVFNLVGV 503
Qy 524 SCETIFLSPGPPFIQNIISYATIGLC--LPFTVILVILCHYKKOFRYESOLOMIO 581
Db 504 SDDTFTVE-----VSKLFTSLIGAAGVUAIFELLVLFLYIKOKPREIRWKIIE 556
Qy 582 VTGPLDNEYFYVDFRDYEDLKWEPFRENLEFGKVLGSGAFGRVMNATAYGISK-TGVSI 640
Db 557 ARE--GNNYTFIDPTQLPVNEKWEPPROKLGKVLGAGAGCKVVEATAFGLGEDKDNFL 614
Qy 641 QVAVKMLKEKADSCKEALMSLKMTHLGHHDNIVNLLGACTLSGPPVLYLIEYCCYGD 700
Db 615 RVAVKMLKANAHSDREALMSLKLHLGHHDNIVNLLGACTYGGPVLVITEYCSLGLD 674
Qy 701 LNYLSKREKERTWETIFKEHNFSSYPTFOAHN-----SSMPGR 742
Db 675 LNFLKQKATFVNLVNI-----PEIMENSNDYKNICNOKWYIRSDSGISSTSS 725
Qy 743 EYOLHPLDQLSGFNGNSHSEDEIYENOKRLAEEEDLNVLTFEDLLCFAYQVAKGM 802
Db 726 YLEMRP-----SQSHTEASGRKSLCEDNGD--WPLDIDDLRLFSLOVAQGL 770
Qy 803 EFLFKSCVHRDLAARNVLVTHGVKVKICDFGLARDILSDSYVYVGNARLPVKWNPES 862
Db 771 DFLASRNCIHRDVAARNVLLTKRAKICDFGLARDIMNDSYVVKGNARLPVKWNPES 830
Qy 863 LFEGLYTKSDVMSYGLLWEIFSLGVNYPGIPVDANFYKLIQSGFKMEQPFYATEGIY 922
Db 831 IFDCVTYVSDVMSYGLLWEIFSLGVNYPGIPVDANFYKLIQSGFKMEQPFYATEGIY 890
Qy 923 FYMQSCWAFDRKRPFPNLTSFLGCOLA--EAEACIRTSIHLPKQAAPQOGRGLRAOS 980
Db 891 MIMKMCWNLPEPTPTFSMISQMINRLLGQDEQELIYRNVO-PEQVAE---GEACDE 945
Qy 981 POR 983
Db 946 PKR 948

RESULT 15
JN0677
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - chicken
N:Alternate names: tyrosine kinase receptor kit
C:Species: Gallus gallus (chicken)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: JN0677
R:Sasaki, E.; Okamura, H.; Chikamune, T.; Kanai, Y.; Watanabe, M.; Naito, M.; Sakurai, M.
Gene 128, 257-261, 1993
A:Title: Cloning and expression of the chicken c-kit proto-oncogene.
A:Reference number: JN0677; MUID:93292995; PMID:7685729
A:Accession: JN0677
A:Molecule type: mRNA
A:Residues: 1-960 <SAS>
A:Cross-references: DDBJ:D13225; NID:g303532; PIDN:BAA02506.1; PID:g303533
A:Experimental source: brain
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;
rotein kinase
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-960/Product: tyrosine kinase receptor #status predicted <MAT>
F:314-380/Domain: immunoglobulin homology <IMM>
F:573-916/Domain: protein kinase homology <KIN>
F:581-589/Region: protein kinase ATP-binding motif
F:76,135,149,269,286,306,318,338,356,453,469/Binding site: carbohydrate (Asn) (covalent)
```

Search completed: May 27, 2003, 14:35:21  
Job time : 32.4866 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:18:49 ; Search time 17.9909 seconds  
(without alignments)  
2286.959 Million cell updates/sec

Title: US-09-919-408-2  
Perfect score: 5264  
Sequence: 1 MRALAQRSDRLLLVLSV.....RGGLRAQSPQVQVTHRS 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5264	100.0	992	1	FLT3_MOUSE
2	4429.5	84.1	992	1	FLT3_HUMAN
3	1258.5	23.9	980	1	KFMS_FELCA
4	1249.5	23.7	978	1	KFMS_FSVMD
5	1223	23.2	975	1	KIT_MOUSE
6	1222.5	23.2	978	1	KIT_CAPHI
7	1221.5	23.2	972	1	KFMS_HUMAN
8	1219	23.2	977	1	KFMS_MOUSE
9	1215	23.1	977	1	KIT_BOVIN
10	1210	23.0	978	1	KFMS_RAT
11	1207	22.9	975	1	KIT_CANFA
12	1196	22.7	976	1	KIT_HUMAN
13	1181	22.4	960	1	KIT_CHICK
14	1178.5	22.4	978	1	KIT_FELCA
15	1157.5	22.0	1088	1	PGDS_RAT
16	1152.5	21.9	1089	1	PGDS_HUMAN
17	1142.5	21.7	1089	1	PGDS_MOUSE
18	1132	21.5	1087	1	PGDS_XENLA
19	1098	20.9	1098	1	PGDR_MOUSE
20	1078	20.5	1106	1	PGDR_HUMAN
21	998.5	19.0	1338	1	VGR1_HUMAN
22	984.5	18.7	1336	1	VGR1_RAT
23	983.5	18.7	1333	1	KIT_FSVHZ
24	967.5	18.4	370	1	KIT_MOUSE
25	965	18.3	1348	1	VGR2_COTJA
26	961.5	18.3	1356	1	VGR2_HUMAN
27	954.5	18.1	1298	1	VGR3_HUMAN
28	952	18.1	1363	1	VGR3_MOUSE
29	934	17.7	1343	1	VGR2_RAT
30	926.5	17.6	1367	1	VGR2_MOUSE
31	818	15.5	813	1	FGR2_XENLA
32	810.5	15.4	823	1	CEK3_CHICK
33	791	15.0	821	1	FGR2_HUMAN

34	789	15.0	806	1	CEK2_CHICK	P18460 gallus gall
35	780	14.8	821	1	FGR2_MOUSE	P21803 mus musculu
36	766.5	14.6	819	1	FGR1_CHICK	P21804 gallus gall
37	761.5	14.5	654	1	BFR2_HUMAN	Q01742 homo sapien
38	758.5	14.4	812	1	FGR1_XENLA	P22182 xenopus lae
39	754.5	14.3	822	1	FGR1_HUMAN	P11362 homo sapien
40	751.5	14.3	806	1	FGR3_HUMAN	P22607 homo sapien
41	751.5	14.3	822	1	FGR1_MOUSE	P16092 mus musculu
42	748.5	14.2	822	1	FGR1_RAT	Q04589 rattus norv
43	732	13.9	801	1	FGR3_MOUSE	Q61851 mus musculu
44	698	13.3	802	1	FGR4_HUMAN	P22455 homo sapien
45	697	13.2	808	1	FGR4_MOUSE	Q03142 mus musculu

ALIGNMENTS

RESULT 1  
ID FLT3\_MOUSE  
AC Q00342; STANDARD; PRT; 992 AA.  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE FL cytokine receptor precursor (EC 2.7.1.112) (Tyrosine-protein kinase  
DE receptor flk-2) (Fetal liver kinase 2) (Tyrosine-protein kinase FLT3).  
GN FLT3 OR FLT-3 OR FLK-2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=91292518; PubMed=1648448;  
RA Matthews W., Jordan C.T., Wiegand G.W., Pardoll D., Lemischka I.R.;  
RT "A receptor tyrosine kinase specific to hematopoietic stem and  
RT progenitor cell-enriched populations.";  
RL Cell 65:1143-1152(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92019834; PubMed=1656368;  
RA Rosnet O., Marchetto S., Delapierre O., Birnbaum D.;  
RT "Murine Flt3, a gene encoding a novel tyrosine kinase receptor of the  
RT PDGFR/CSF1R family.";  
RL Oncogene 6:1641-1650(1991).  
RN [3]  
RP CHARACTERIZATION.  
RX MEDLINE=93205405; PubMed=8384358;  
RA Maroc N., Rottapel R., Rosnet O., Marchetto S., Lavezzi C.,  
RA Mannoni P., Birnbaum D., Dubreuil P.;  
RT "Biochemical characterization and analysis of the transforming  
RT potential of the FLT3/FLK2 receptor tyrosine kinase.";  
RL Oncogene 8:909-918(1993).  
CC -!- FUNCTION: RECEPTOR FOR THE FL CYTOKINE. HAS A TYROSINE-PROTEIN  
CC KINASE ACTIVITY.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: HEMATOPOIETIC STEM AND PROGENITOR CELL-  
CC ENRICHED POPULATIONS. FOUND IN BRAIN, PLACENTA AND TESTIS.  
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
CC PROTEIN KINASES.  
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----



DR EMBL; M64689; AAA37634.1; -;  
 DR EMBL; X59398; CAA42041.1; -;  
 DR PIR; A39931; A39931.  
 DR HSP; P11362; IFGK.  
 DR MGD; MGI:95559; Flt3.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003600; Ig\_Like.  
 DR InterPro; IPR001824; RTKinaseIII.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00047; ig; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 2.  
 DR SMART; SM00410; IG\_Like; 1.  
 DR SMART; SM00219; TyrcK; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR Signal; Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;  
 KW Glycoprotein; Phosphorylation; ATP-binding; Immunoglobulin domain.  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 992 FL CYTOKINE RECEPTOR.  
 FT DOMAIN 28 544 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 545 564 POTENTIAL.  
 FT DOMAIN 565 992 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 611 946 PROTEIN KINASE.  
 FT NP\_BIND 617 625 ATP (BY SIMILARITY).  
 FT BINDING 645 645 ATP (BY SIMILARITY).  
 FT ACT\_SITE 814 814 BY SIMILARITY.  
 FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 474 474 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 503 503 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 150 150 R -> A (IN REF. 2).  
 FT CONFLICT 242 242 C -> S (IN REF. 2).  
 FT CONFLICT 726 726 S -> F (IN REF. 2).  
 FT CONFLICT 957 979 CIRTSIHLPKQAAPOQRGGLRAQ -> MYQNMGGNVPEHPS  
 FT IYQNRRLPSREAGSEPP (IN REF. 2).  
 FT CONFLICT 983 983 R -> A (IN REF. 2).  
 FT SEQUENCE 992 AA; 112639 MB; 407A087853372100 CRC64;

Query Match 100.0%; Score 5264; DB 1; Length 992;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 992; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRALQSRDRLLLVLSVMIETVNTQDLPVTKVLISHENNGSSAGKPSRYMRVGRS 60  
 DB 1 MRALQSRDRLLLVLSVMIETVNTQDLPVTKVLISHENNGSSAGKPSRYMRVGRS 60  
 QY 61 PEDLOCTPRROSEGVVEAATVEAEGSITLQVLAATPGDLSCLWFKHSLGCOPHFD 120  
 DB 61 PEDLOCTPRROSEGVVEAATVEAEGSITLQVLAATPGDLSCLWFKHSLGCOPHFD 120  
 QY 121 LQNGRIVSMALNTVTOAGEYLLHIOSEERANYTVLFTVNVYRDQYVLRPRYFRKMNQ 180  
 DB 121 LQNGRIVSMALNTVTOAGEYLLHIOSEERANYTVLFTVNVYRDQYVLRPRYFRKMNQ 180  
 QY 181 DALLCISEGVPEPTVWVLCSSHRESCKEKGPAVVRKEEYVHLEFGDTRCCARNALGR 240  
 DB 181 DALLCISEGVPEPTVWVLCSSHRESCKEKGPAVVRKEEYVHLEFGDTRCCARNALGR 240  
 QY 241 ECTKLFTIDLNQAPQSTLPQLFLKVGEPWLWIRKAIHNVHGFGLTWELEDKALEGGSYFE 300  
 DB 241 ECTKLFTIDLNQAPQSTLPQLFLKVGEPWLWIRKAIHNVHGFGLTWELEDKALEGGSYFE 300  
 QY 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSSKHPQSALVTILEKGFINATSSQEEY 360

DB 301 MSTYSTNRTMIRILLAFVSSVGRNDTGYTCCSSSKHPQSALVTILEKGFINATSSQEEY 360  
 QY 361 EIDPYEKFCSVRFKAYPRICTWIFSOASPCPEORGLGDSYISKFCDDHKNKPEYIFY 420  
 DB 361 EIDPYEKFCSVRFKAYPRICTWIFSOASPCPEORGLGDSYISKFCDDHKNKPEYIFY 420  
 QY 421 AENDDAOFTKMTLNIRKKPOVLANASQASCSGSDGYPLPSWTWKKCDKSPNCTEETIP 480  
 DB 421 AENDDAOFTKMTLNIRKKPOVLANASQASCSGSDGYPLPSWTWKKCDKSPNCTEETIP 480  
 QY 481 EGVNKKANRKFVGOVSSSTLNSEAGKGLLVKCCANSMGTSCETIFLNSGGPFPEIQ 540  
 DB 481 EGVNKKANRKFVGOVSSSTLNSEAGKGLLVKCCANSMGTSCETIFLNSGGPFPEIQ 540  
 QY 541 DNISFYATIGLCPLFIVVLIVLICHYKQKQFYESQLOMIQVTPGLDNEYFYVDFRDY 600  
 DB 541 DNISFYATIGLCPLFIVVLIVLICHYKQKQFYESQLOMIQVTPGLDNEYFYVDFRDY 600  
 QY 601 DLKWEFFRENLEFGKVLGSGAFGRVMNATAYGISKTGYSIQVAVKMLKEKADSCKEALM 660  
 DB 601 DLKWEFFRENLEFGKVLGSGAFGRVMNATAYGISKTGYSIQVAVKMLKEKADSCKEALM 660  
 QY 661 SELKMMTHLGHHDNIVNLLGACTLSGPVYLFEYCCYGDLLNLYLSKREKFKHRTWTEIEFK 720  
 DB 661 SELKMMTHLGHHDNIVNLLGACTLSGPVYLFEYCCYGDLLNLYLSKREKFKHRTWTEIEFK 720  
 QY 721 EHNFSSTYPTFOAHNSNMPGSGREVOLHPDLQSLGFNGNSIHSEDEIEYENOKRLAESEE 780  
 DB 721 EHNFSSTYPTFOAHNSNMPGSGREVOLHPDLQSLGFNGNSIHSEDEIEYENOKRLAESEE 780  
 QY 781 EDLVNLTPEDDLFCFAYQVAKGMFELEKSCVHRDLAARNVLVTHGKVKVVICDFGLARDIL 840  
 DB 781 EDLVNLTPEDDLFCFAYQVAKGMFELEKSCVHRDLAARNVLVTHGKVKVVICDFGLARDIL 840  
 QY 841 SDSSVYVVRGNARLPVKWMAPELSFEGITYIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900  
 DB 841 SDSSVYVVRGNARLPVKWMAPELSFEGITYIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900  
 QY 901 FYKLQSGFKMEQPPYATGEGYFVMSQWAFDSRRKRPFPNLTSLGCOLAEAEACIRT 960  
 DB 901 FYKLQSGFKMEQPPYATGEGYFVMSQWAFDSRRKRPFPNLTSLGCOLAEAEACIRT 960  
 QY 961 SIHLPKQAAPOQRGGLRAQSPQKVKIHRERS 992  
 DB 961 SIHLPKQAAPOQRGGLRAQSPQKVKIHRERS 992

RESULT 2  
 FLT3\_HUMAN  
 ID FLT3\_HUMAN STANDARD; PRT; 993 AA.  
 AC P36888; O13414;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE FL cytokine receptor precursor (EC 2.7.1.112) (Tyrosine-protein kinase receptor FLT3) (Stem cell tyrosine kinase 1) (STK-1) (CD135 antigen).  
 GN FLT3 OR STK1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=94119906; PubMed=7507245;  
 RA Small D., Levenstein M., Kim E., Carow C., Amin S., Rockwell P., Witte L., Burrow C., Ratajczak M.Z., Gewirtz A.M., Civin C.I.;  
 RA "STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in CD34+ human bone marrow cells and is involved in the proliferation of early progenitor/stem cells.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:459-463(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.





```
Db 522 INALLLLLLLLLYKKYKQPKYQVRKKIIE--SYEGNSYTFIDPTQLPYNEKWEFFRNLL 579
QY 612 EFGKVLGSAFGVMNATAYGISTQVAVKMLKEKADSCKEKALMSKMMTHLGH 671
Db 580 QFGKTLGAGAFKGVVEATAFGLGKEDAVLKVAVKMLKSTAHADKEKALMSKIMSHLGQ 639
QY 672 HENIVNLGACLTSGPVYLIFECYCCGDLNLYLSKRE-----KF 711
Db 640 HENIVNLGACHTGGPVLVITEYCCYCGDLNLFURQAQEAAMLGSPSLVSGODPEAGAYKNI 699
QY 712 HRTWTBIFKEHNFSSYPTFOAHSSNMPGSRREVQLHPPLDQLSGNGNSHSEDEIYEN 771
Db 700 HLEKKYVRDSDES-----SOGVDYIVEMRPVSTSSNDSFSEDL---- 740
QY 772 OKRLAEEEDLNVLFEDLLCFAYQVAKGMEFLPKSCVHRDLAARNVLVTHGKVKIC 831
Db 741 -----GREGRPLELRDLHLHFSQVAGMAFLASKNCIHRDVAARNVLLTSGRVAKIG 793
QY 832 DFLGARDILSDSSVYVGRNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIIFSLGVNP 891
Db 794 DFLGARDIMDNSYIVKGNARLPVKWMAPESPFDCVTVQSDVWSYGILLWEIIFSLGLNP 853
QY 892 YGICPVDAFYKLIQSGFKMEOPFVATEGIYFVMQSCWAFDSKRPSPFNLTSTFLGCOLA 951
Db 854 YGICLVNSFYKLVKDYQMAOPAFAPAKNLYISINQAWALEPRTTRTFQOICSL--QKQ 911
QY 952 EAEACIRTSIHLPKOAPOQ-----RGLRAQSPQRQ 984
Db 912 AOEDRRVPVNTLPPSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 953

RESULT 4
KFMS.FSVM
AC P00545: 086597; STANDARD; PRT; 978 AA.
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase transforming protein fms (EC 2.7.1.112).
GN V-FMS.
OS Feline sarcoma virus (strain McDonough).
OC Viruses; Retroid viruses; Retroviridae; Mammalian type C retroviruses.
OX NCBI_TaxID=11778;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84119469; PubMed=6582485;
RA Hampe A., Gobet M., Sherr C.J., Galibert F.;
RT "Nucleotide sequence of the feline retroviral oncogene v-fms shows
RT unexpected homology with oncogenes encoding tyrosine-specific protein
RT kinases.";
RL Proc. Natl. Acad. Sci. U.S.A. 81:85-89(1984).
RN [2]
RP REVISIONS, SEQUENCE FROM N.A.
RX MEDLINE=92015516; PubMed=1833563;
RA Smola U., Hennig D., Hadwiger-Fangmeier A., Schuetz B., Pfaff E.,
RA Niemann H., Tamura T.;
RT "Reassessment of the v-fms sequence: threonine phosphorylation of the
RT COOH-terminal domain.";
RL J. Virol. 65:6181-6187(1991).
CC -!- FUNCTION: V-FMS IS DERIVED FROM THE RECEPTOR FOR COLONY
CC STIMULATING FACTOR 1 (CSF-1).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-FMS
CC POLYPROTEIN.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC
```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: K01643; AAA3045.1; -  
PIR: A0654; TVNMD.  
HSP: P11362; IFGK.  
InterPro: IPR00719; Euk\_pkinase.  
InterPro: IPR003006; Ig\_MHC.  
InterPro: IPR003598; Ig\_c2.  
InterPro: IPR003600; Ig\_like.  
InterPro: IPR001824; RTKinaseIII.  
InterPro: IPR001245; Tyr\_kinase.  
Pfam: PF00047; ig; 3.  
Pfam: PF00069; pkinase; 1.  
ProDom: PD000001; Euk\_pkinase; 2.  
SMART: SM00410; IG\_like; 4.  
SMART: SM00408; IGC2; 1.  
SMART: SM00219; TyrKc; 1.  
PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
PROSITE: PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
PolyProtein: Tyrosine-protein kinase; Oncogene; Transferase; Receptor;  
KW Transmembrane; Glycoprotein; Phosphorylation; ATP-binding;  
KW Immunoglobulin domain; Repeat.  
FT DOMAIN 1 543  
FT TRANSMEM 544 568  
FT DOMAIN 569 978  
FT DOMAIN 58 138  
FT DOMAIN 141 231  
FT DOMAIN 238 332  
FT DOMAIN 333 431  
FT DOMAIN 432 536  
FT DOMAIN 613 942  
FT NP\_BIND 619 627  
FT BINDING 647 647  
FT DISULFID 76 118  
FT DISULFID 161 211  
FT DISULFID 258 312  
FT DISULFID 451 516  
FT MOD\_RES 841 841  
FT MOD\_RES 973 973  
FT CARBOHYD 79 79  
FT CARBOHYD 107 107  
FT CARBOHYD 128 128  
FT CARBOHYD 187 187  
FT CARBOHYD 309 309  
FT CARBOHYD 320 320  
FT CARBOHYD 336 336  
FT CARBOHYD 369 369  
FT CARBOHYD 444 444  
FT CARBOHYD 511 511  
FT CARBOHYD 524 524  
FT CONFLICT 714 714  
FT CONFLICT 971 978  
FT SEQUENCE 978 AA; 108491 MW; 4C7CAC4835185EBF CRC64;  
Query Match 23.7%; Score 1249.5; DB 1; Length 978;  
Best Local Similarity 33.2%; Pred. No. 3.9e-76;  
Matches 328; Conservative 162; Mismatches 348; Indels 149; Gaps 32;  
QY 75 TVTEATVEVAESGITVLQVQLATPGDLSCV-----W---VFKHSSLCGCPHDLQNRG 125  
DB 47 TAWHAQGVPIQPSGPELVPEPTVTLRCVNGSVEMDGPISPHWNLDLPSSSI---- 102  
QY 126 IVSMAILNVETQAGEYLLHI---QSERANYTVLFTVNVVRDTQLYVL--RRPY----- 173  
DB 103 ---LTTNNATFQNTGT--HCTEPGNRGGNATI-----HLVVKDPRPWKVLQAE 148

QY 174 FRKMNODALL-CISEGVPEPTVWVLCSSHRECKEFGPAVVRKDEKVLHFGTDI-- 230  
 Db 149 VTVEGQDALLPCL---LTDPALE-----AGVSLVRVRGPPVLRQTNYSFSPWHGTIHK 200  
 QY 231 -----RCCARNALGRECTKL-----FTIDLNOAPOSTL-PQLPLKV-GEPLWIRCK 274  
 Db 201 AKFIENHVQCSAR-VDGRVTVSMGIWLKVQKDISGPATLLEPAELVRIQGEAAQIVCS 259  
 QY 275 AIVNHGFGLTWELEKALEEGSYFEMSYSTNRTMIRILLAFVSSVGRNDTGYTCSSS 334  
 Db 260 ASNIDNVFVSLRHGDTKL---TISQOSDFHNRQ- KVLTLNLDHVSFODAGNYSCTAT 315  
 QY 335 K---HPSQALVTRILEKFINATSSOE-EYEIDPYKFCFSVRKAYPRRC-TWI----- 385  
 Db 316 NAWGNHSASWFRVESAIVNSLTSEOSLQEVTVGEKVDLOVKEAYPGLESEFNWYVLP 375  
 QY 386 FSOASPPCBOGLEDDGYISKFCD----HKNKGPEYIFVAENDDAOFTKMTNIRKKPQ 441  
 Db 376 FSDYQKDLDFVTIKDTRYRTSTLSPLKRSSESGYSFLARNAGGQNALTFELTLRYPE 435  
 QY 442 VIANASASQAS-----CSSDGYPLPSWTKKCDKSPNCTEE---IPEGVWKKANRKFV 494  
 Db 436 VRVTMLINGSDFLLCEASGYPOPSVTWQCRSHTRDCDESAGLVLEDSHSEVLSPVPFY 495  
 QY 495 QWYSSSTLNSEAGKGLLVKCCAYNSMTSCETIFLNSPGFPFFFIODNIFATIGLC-- 552  
 Db 496 EVIVHSLAIGTLEHRTYECRAFNSVGNSSQTFWPLISGAHTPLDPLDLFTPVLLTMS 555  
 QY 553 -LPFIVVLIVLICHKKYKQRYEQSLOMIOVTGPDNEXEYVDFRDYEDLKWFEPPREN 611  
 Db 556 IMALLLLLLLLLYKKYKQPKYQVRWKIIE--SYEGNSYTFIDPTQLPYNEKWEFFPRNL 613  
 QY 612 EPKVLGSGAFGRVMNATAGISKTGVSIOVAVKMLKEKADSEKALMSLKMMTHLGH 671  
 Db 614 QFGKTLTGAFGKVEATAFGLGKEDAVLVAVKMLKSTAHADSEKALMSLKIMSHLQ 673  
 QY 672 HDNIVNLGACTLGGPVLYFECYGDLLNLYRSKRE-----KF 711  
 Db 674 HENIVNLGACTHGGPVLYFECYGDLLNLYRSKRE-----KF 711  
 QY 712 HRTWTEIFKHEHNSYPTFOAHNSMSPGREGVQLHPPDLQSLGFGNSIHSDEIEYEN 771  
 Db 734 HLEKVVYRRDSGFS-----SOGVDYVEMRPVSTSSNDGSFSEEDL----- 774  
 QY 772 OKRLAEEEDLNLVTFEDLLCFAYQVAKMEFLFECFVSRDLAARNLVTHCKVVKIC 831  
 Db 775 -----GKEDGRPLELDLHFSQVQAQMAFLASKNCTRDRVAARNVLLTSGRVAKIG 827  
 QY 832 DFLGARDLSDSVVVRGNARLPVKWMAPELFGIYTIKSDVMSYGILLWEIFSLGNP 891  
 Db 828 DFLGARDIMDSNVIVKGNARLPVKWMAPELFGIYTIKSDVMSYGILLWEIFSLGNP 887  
 QY 892 YPGIVPDANFYKLQSGFKMEOPFYATEGIVFMOSCAFDSRKRPSFPNLTSLFGQLA 951  
 Db 888 YPGILVNSKFKLVKQGYQMAQAPAFKNIYSIMOACWALEPRTTRPTFQICSL--OKQ 945  
 QY 952 EAEACIITSILHPKQAA-----PQOR 973  
 Db 946 AODRRVPNTNLFSSSSSRLLRPWR 972

RESULT 5  
 KIT\_MOUSE  
 ID KIT\_MOUSE STANDARD; PRT; 975 AA.  
 AC P05532; Q61415; Q61416; Q61417;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Mast/system cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)  
 DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).  
 GN KIT OR SL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/C;  
 RX MEDLINE=88296403; PubMed=2456920;  
 RA Qiu F., Ray P., Brown K., Barker P.E., Jhanwar S., Ruddle F.H.,  
 RA Besmer P.;  
 RT "Primary structure of c-kit: relationship with the CSF-1/PDGF  
 RT receptor kinase family -- oncogenic activation of v-kit involves  
 RT deletion of extracellular domain and C terminus.";  
 RL EMBO J. 7:1003-1011(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A. (TRUNCATED FORM).  
 RC STRAIN=ICR;  
 RX MEDLINE=92331813; PubMed=1378413;  
 RA Rossi P., Marzitali G., Albanesi C., Charlesworth A., Geremia R.,  
 RA Sorrentino V.;  
 RT "A novel c-kit transcript, potentially encoding a truncated receptor,  
 RT originates within a kit gene intron in mouse spermatids.";  
 RL Dev. Biol. 152:203-207(1992).  
 RN [3]  
 RP LIGAND.  
 RX MEDLINE=91006023; PubMed=1698611;  
 RA Tan J.C., Buck J., Levi E., Besmer P.;  
 RT "Candidate ligand for the c-kit transmembrane kinase receptor: KL, a  
 RT fibroblast derived growth factor stimulates mast cells and erythroid  
 RT progenitors.";  
 RL EMBO J. 9:3287-3294(1990).  
 RN [4]  
 RP VARIANT W42 ASN-790.  
 RX MEDLINE=90100577; PubMed=1688471;  
 RA Tan J.C., Nocka K., Ray P., Traktman P., Besmer P.;  
 RT "The dominant W42 spotting phenotype results from a missense mutation  
 RT in the c-kit receptor kinase.";  
 RL Science 247:209-212(1990).  
 RN [5]  
 RP VARIANTS W37 LYS-582; W41 MET-660 AND W41 MET-831.  
 RX MEDLINE=90269214; PubMed=1693331;  
 RA Nocka K., Tan J.C., Chiu E., Chu T.Y., Ray P., Traktman P.,  
 RA Besmer P.;  
 RT "Molecular bases of dominant negative and loss of function mutations  
 RT at the murine c-kit/white spotting locus: W37, W41 and W.";  
 RL EMBO J. 9:1805-1813(1990).  
 CC -I- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL  
 CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.  
 CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND  
 CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL  
 CC 3-KINASE (PI3K).  
 CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -I- DISEASE: WHITE-SPOTTING VARIANTS INDUCES SEVERE EFFECTS ON  
 CC PIGMENTATION, GEMETOGENESIS AND HEMATOPOIESIS. MICE HOMOZYGOUS  
 CC FOR W42 DIE PERINATALLY OF MACROCYTIC ANEMIA.  
 CC -I- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -I- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 CC  
 CC EMBL; Y00864; CAA68772.1;  
 DR EMBL; X65997; CAA46798.1;  
 DR EMBL; X65998; CAA46799.1; ALT\_SEQ.  
 DR EMBL; X65998; CAA46800.1;  
 DR PIR; S00474; TVMSKT.  
 DR HSSP; P11362; IFGK.

DR MGI:96677; Kit.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003598; Ig\_C2.  
 DR InterPro: IPR003600; Ig\_Like.  
 DR InterPro: IPR001824; RtkinaseII.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00047; Ig\_1.  
 DR Pfam: PF00069; pkinase; 1.  
 DR ProDom: PD000001; Euk\_pkinase; 2.  
 DR SMART: SM00410; IG\_Like; 3.  
 DR SMART: SM00408; IGC2; 1.  
 DR SMART: SM00219; TyKc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;  
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;  
 KW Immunoglobulin domain; Disease mutation.  
 FT SIGNAL 1 22  
 FT CHAIN 23 975 MAST/STEM CELL GROWTH FACTOR RECEPTOR.  
 FT DOMAIN 23 519 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 520 542 POTENTIAL.  
 FT DOMAIN 543 975 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 588 935 PROTEIN KINASE.  
 FT NP\_BIND 594 602 ATP (BY SIMILARITY).  
 FT BINDING 622 622 ATP (BY SIMILARITY).  
 FT ACT\_SITE 790 790 BY SIMILARITY.  
 FT MOD\_RES 821 821 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 146 146 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 303 303 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 323 323 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 355 355 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 370 370 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 466 466 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 489 489 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARIANT 1 12 MGVAGAWDLIC -> MAVAVFPFLPQQ (IN TRUNCATED FORM).  
 FT VARIANT 13 785 MISSING (IN TRUNCATED FORM).  
 FT VARIANT 582 582 E -> K (IN W37 SPOTTING).  
 FT VARIANT 660 660 T -> M (IN W4 SPOTTING).  
 FT VARIANT 790 790 D -> N (IN W42 SPOTTING).  
 FT VARIANT 831 831 V -> M (IN W41 SPOTTING).  
 FT SEQUENCE 975 AA; 109001 MW; BAB5CA4D9AF9CD2A CRC64;  
 Query Match 23.2%; Score 1223; DB 1; Length 975;  
 Best Local Similarity 32.1%; Pred. No. 2.3e-74;  
 Matches 324; Conservative 172; Mismatches 331; Indels 182; Gaps 39;  
 QY 36 CVLISHENNGSSAGKPSRYRMVRGSPEDLOCTPRROSEGTVYEAAATVEVAESGSIITLQVQ 95  
 DB 12 CVLLVLLRGOTATNSQPSA-----SPCEPSPPIHQAQSELIVEAGD-TL--- 54  
 QY 96 LATPGLDLSL-----WVFHSSLGCPHF-----DLQNRGIVSMALNVTETQAGVYLL 144  
 DB 55 -----SLTCTDIDPQVVRWTFK-----TYFNEMVENKKNWIOEKA-----EATRTGYTC 98  
 QY 145 HIOSEARNTVLTIVNVRD-TOLYVLRPPYFRKMNODALL-----CIS 187  
 DB 99 ---SNSNGLTSSITVYFVRPAKFLVLGLPLFGK-EDSDALVRCPLTDPPQVSNVSLIEDCG 154  
 QY 188 EG-----VPEPTVEVWLCSSHRESCKEKGPAVVRKEKVLHFLCTDIRCCARNALGR 240  
 DB 155 KSLPTDLTFTFNP-----KAGITIKNVKRAYH-----RLCVRCAQR 191  
 QY 241 ECKTL-----FTIDLNQ-----POSTLPQL--FLKVGPLMIRCKAIHNVHFGTLW--- 286  
 DB 192 DGTWHLSDKFTLKVREAIKAIPVVSVPETSHLLKKGDTFTVVCITKIDYSTVSNMMLKMN 251  
 QY 287 -----ELEDKALEGSEYFEMSTYNRTNRMIRILLAFVSSVGRNDTGYTCSSSKRPSQ 339  
 DB 339 -----ELEDKALEGSEYFEMSTYNRTNRMIRILLAFVSSVGRNDTGYTCSSSKRPSQ 339

DB 252 POQHIAQVKNHNSWHRGDF-----NYERQETLT-----ISSARVDDSGVFMCYANNTFGS 301  
 QY 340 SALVT---ILEKGFNATS-SQEEYEDPYEKFCFSVRFKAYPR-IRCTWIFSASFPCE 394  
 DB 302 ANVTYTTLLKVVKEGFINISPVKNTTFVTDGENVLDVVEYEAYPKPEHQOQWYIMRT--SA 359  
 QY 395 ORGLE-----DGYSTISKCDH-----KNKPGVEYIFAENDDAOFTKMTINIRKKKPOVL 443  
 DB 360 NKGKDYKSDNKSINIRVYNQLRLRLKGTGGYTFVLVNSDASASVTFFVYVNTKPEIL 419  
 QY 444 A--NASASQASCSGSGYPLPSWTWKCKSDKSPNCTEBI-PEGVNNKANRKFVQWVSSS 500  
 DB 420 TYDRLINGMLQCVAEQPEPTIDWYFCTGAEQRCTTPVSPVDVQVQVNSVSPFGKLVVQS 479  
 QY 501 TLNMSKAGLLVKCCAYNSMTGSCETIFLNSPPGPFQIDNISFYATIGLCLFIV--- 557  
 DB 480 SIDSVFRHNGTVECKASNDVGKS--SAFFN---FAFKEQIOAHTLFTPLLIGFVVAAG 533  
 QY 558 ---VLIVLICHKVKQFRIESQLQMI-QVTGPLDNEYFYVDFRDEYDLDKWEFFPRENLEF 613  
 DB 534 AMGLIVNVLTYKYLQKPMYEQWKVVEING---NNVYIDPTOLPYDHKWEFFPRNRLSF 590  
 QY 614 GKVLGSGAFGRVMNATAYGISTGVSIQVAVKMLKEKADSCKEKALMSKLMTHLGHHD 673  
 DB 591 KGTLGAGAFGVVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSKLVSLGNHM 650  
 QY 674 NIVNLLGACTLSPGVYLIFEYCCYGDLLNLYRSKREX-----HRTWTEIFKHNFSY 727  
 DB 651 NIVNLLGACTVGGTLVITEYCCYGDLLNLYRSKREX-----HRTWTEIFKHNFSY 710  
 QY 728 PTQAHNSNM---PGSREVQLHPDQLSFGNCSIHSEDEIEYENOKRLAEEDLN 784  
 DB 711 PSCDS-SNEYMDMKPGVSYV-VPTKTDK-----RRSARIDSYIERDVTTPAIMEDELALD 763  
 QY 785 VLAFEDLCEPAYQVAKGMEFLEEKSCVHRDLAARNVLVTHGKVKVKICDFGLARDLSDSS 844  
 DB 764 L---DILLSFSYQAKAMAFASKNCHIRDLAARNILLTHGRITKICDFGLARDLRNSN 820  
 QY 845 YVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPGDANFYKL 904  
 DB 821 YVVGNGARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPGDANFYKL 880  
 QY 905 IQSGFKMEQPFYATEGIYFVWQSCWAFDSRKRPSFNPNTSLGCOLAEA 953  
 DB 881 IKGFRMVSPHAPAEYDMVKTCWDADPLKRPFTKQVQVLIEXQISDS 929  
 RESULT 6  
 KIT\_CAPHI  
 ID KIT\_CAPHI STANDARD; PRT; 978 AA.  
 AC Q28317;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)  
 DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).  
 GN KIT.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Capra.  
 OX NCBI\_TaxID:9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Shiba: TISSUE=Cerebellum;  
 RX MEDLINE=97342548; PubMed=9195245;  
 RA Tanaka S., Yanagisawa N., Tojo H., Kim Y.-J., Tsujimura T.,  
 RA Kitamura Y., Sawasaki T., Tachi C.;  
 RT "Molecular cloning of cDNA encoding the c-kit receptor of Shiba goats  
 RT and a novel alanine insertion specific to goats and sheep in the  
 RT kinase insert region.";  
 RL Blochim. Biophys. Acta 1352:151-155(1997).  
 CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL

[illegible]





```
Db 172 SDYQCSALMG-GR---KVMISIRLKVKQKVIPIGPPALTLPVPAELVIRGEAAQIVCSAS 227
Qy 277 HVNNGFGLTWLEOKALEEGSYFEMSTYSTRNTRILLAFVSSVGRNDGTGYTCSSS-- 334
Db 228 SVDNFDVFLQHNNTKL---AIPOOSDHNRYO-KVLTLLDQVDFOHAGNYSVANSV 283
Qy 335 --KHPSOSALVTILEKGFINTSQE-EYEIDPEKFCFSVRFRKAYPRIR-CTWIFSOAS 390
Db 284 QGKH-STSMFFRVVESAYNLNLSSEQLNLOEVTGEGNLKVMVEAYPGLGFNWTY---- 338
Qy 391 FPCGORGLEDGYSIKFCDHNKP-----GEYIFYAEN 423
Db 339 -----LGPFSHDQPEPKLANATTKDYRHTFTLSLPLKPKSEAGRYSFLARN 385
Qy 424 DDAQFTKMTNIRKKPOV-----LANASASQSSDGYPLPSPWTWKKCKDSKSPNCTEE 478
Db 386 PGGWRALTFTLRYPPVESVITWTFINGSGT-LLCAASGYPOPNVTWLCQSGHTORCDEA 444
Qy 479 IPEGYN---KANKRVFGOWSVSSTLNMSEAGKGLLVKCCAYNMGTSCTETIFLNSPG 534
Db 445 QVLQWDDDPYEPVLSQEPFKVTVQSLTVEPLEHNTQYECRAHNSVSGSWAFIPISAG 504
Qy 535 PPFPTQDNISFYATGLC---LPFIWLVILCHYKKQFRYESOLOMTQVTPGLDNEYF 591
Db 505 ANTHPPDEFLTPTVVVACMSINALLLLLLLLYKQKPKYQVRWKIIE--SYEGNSYT 562
Qy 592 YVDFRDYEDLKWEFPRENLEFGKLGSGAGFGRVNMATAYGSKTGVSTQVAVKMLKEKA 651
Db 563 FIDPTOLPYNEKWEPPRNLLQFKGLGAGAFKVVVEATFGLGKEDAVLKVAVKMLKSTA 622
Qy 652 DSCKEALMSKLMTHLGHHDNIYNLLGACTLSGPVYLIFBYCCYCGDLLNLYLRSKREKF 711
Db 623 HADEKALMSKLMKSHLQGHENIVNLLGACTHGGPVLVITEVCCYCGDLLNLYLRKAE-- 680
Qy 712 HRTWTEIFEKHNFSYPTPTQAHNSNMPGSRVQLPDLQSLGFGNGNSIHSEDE----- 766
Db 681 -----AMLGP---SLSPGQDPEGGVYKNIHLEKKYVRD 712
Qy 767 -----IEYENQKLAEE--EEDLNVLTFEDLLCFAYQVAKGMFLEFK 808
Db 713 SGFSSGVDTYVEMRPVSYSSNDSSEQLDKEDGRLPLERLDLLHFSQVQAQMAFLASK 772
Qy 809 SCVHRDLAARNVLVTHGKVKVICDFGLARDILSDSSVYVYVGNARLPVKNWAPESLFEIGY 868
Db 773 NCIHVRDAARNVLLTHGKVKVICDFGLARDINNSVYVYVGNARLPVKNWAPESIEDCVY 832
Qy 869 TKSDVMSYGILLWEITSLGVPNDYPIGVNDANFYKLIQSGFKWEQFPYATEGIFYWQSC 928
Db 833 TVQSDVMSYGILLWEITSLGVPNDYPIGVNDANFYKLIQSGFKWEQFPYATEGIFYWQSC 892
Qy 929 WAFDSKRPSFPNLTSLGCOLAAEEACIRTSIHLPKQAPQOGRGLRAQSPQ 982
Db 893 WALEPHTPTFOQICSF--QEOAQEDRRERYTNLPSSS---RSGSGSSSSE 941

RESULT 8
KFMS_MOUSE
ID KFMS_MOUSE STANDARD; PRT: 977 AA.
AC P09581; Q9DBH9;
DT 01-MAR-1989 (Rel. 10, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)
DE (EC 2.7.1.12) (Fms proto-oncogene) (c-fms).
GN CSF1R OR CSFMR OR FMS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88217329; PubMed=2966922;
Rothwell V.M., Rohrschneider L.R.;
```

```
RT "Murine c-fms cDNA: cloning, sequence analysis and retroviral
RT expression.";
RL Oncogene Res. 1:311-324(1987).
RN [2]
RP Rothwell V.M.;
RL Submitted (SEP-1988) to the EMBL/GenBank/DBJ databases.
RN [3]
RP REVISIONS.
RX MEDLINE=93181280; PubMed=8441691;
RA de Parseval N., Boreaux D., Gisselbrecht S., Sola B.;
RT "Reassessment of the murine c-fms proto-oncogene sequence.";
RL Nucleic Acids Res. 21:750-750(1993).
RN [4]
RP SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=21083660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubil F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [5]
RP SEQUENCE OF 1-16 FROM N.A.
RX MEDLINE=93268269; PubMed=8497248;
RA Yue X., Favot P., Dunn T.L., Cassidy A.I., Hume D.A.;
RT "Expression of mRNA encoding the macrophage colony-stimulating factor
RT receptor (c-fms) is controlled by a constitutive promoter and tissue-
RT specific transcription elongation.";
RL Mol. Cell. Biol. 13:3191-3201(1993).
RN [6]
RP AUTOPHOSPHORYLATION SITES.
RX MEDLINE=90258890; PubMed=2160591;
RA van der Geer P., Hunter T.;
RT "Identification of tyrosine 706 in the kinase insert as the major
RT colony-stimulating factor 1 (CSF-1)-stimulated autophosphorylation
RT site in the CSF-1 receptor in a murine macrophage cell line.";
RL Mol. Cell. Biol. 10:2991-3002(1990).
CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN
CC TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X06368; CAA29666.1; ALT_SEQ.
DR EMBL; AK004947; BAB23691.1; -.
DR EMBL; S62219; -. NOT_ANNOTATED_CDS.
```



RX MEDLINE=94215924; PubMed=7512939;  
 RA Kubota T., Hikono H., Sasaki E., Sakurai M.;  
 RT "Sequence of a bovine c-kit proto-oncogene cDNA.";  
 RL Gene 141:305-306(1994).  
 CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL  
 CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.  
 CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND  
 CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL  
 CC 3-KINASE (PI3K).  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; D16680; BAA04084.1; -  
 CC HSSP; P11362; LFGL.  
 CC InterPro: IPR000719; Euk\_pkinase.  
 CC InterPro: IPR003006; Ig\_MHC.  
 CC InterPro: IPR003598; Ig\_c2.  
 CC InterPro: IPR003600; Ig\_like.  
 CC InterPro: IPR001824; RTKinaseIII.  
 CC InterPro: IPR001245; Tyr\_pkinase.  
 CC Pfam; PF00047; Ig; 1.  
 CC Pfam; PF00069; pkinase; 1.  
 CC Pfam; PD000001; Euk\_pkinase; 2.  
 CC SMART; SM00410; IG\_like; 2.  
 CC SMART; SM00408; IGC2; 1.  
 CC SMART; SM00219; TyrcK; 1.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 CC PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 CC Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;  
 CC Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;  
 CC Immunoglobulin domain.  
 FT SIGNAL 1 22  
 FT CHAIN 23 977  
 FT MAST/STEM CELL GROWTH FACTOR RECEPTOR.  
 FT DOMAIN 23 521  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 522 544  
 FT POTENTIAL.  
 FT DOMAIN 545 977  
 FT CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 590 938  
 FT PROTEIN KINASE.  
 FT NP\_BIND 596 604  
 FT ATP (BY SIMILARITY).  
 FT BINDING 624 624  
 FT ATP (BY SIMILARITY).  
 FT ACT\_SITE 793 793  
 FT BY SIMILARITY.  
 FT MOD\_RES 824 824  
 FT PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 94 94  
 FT N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 130 130  
 FT N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 145 145  
 FT N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 284 284  
 FT N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 294 294  
 FT N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 301 301  
 FT N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 321 321  
 FT N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 353 353  
 FT N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 368 368  
 FT N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 401 401  
 FT N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 464 464  
 FT N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 487 487  
 FT N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 487 487  
 FT N-LINKED (GLCNAC...) (POTENTIAL).  
 FT SEQUENCE 977 AA; 109685 MW; 4B2719050883B7EF CRC64;

Query Match

23.1%; Score 1215; DB 1; Length 977;

Best Local Similarity 32.8%; Pred. No. 8.1e-74;

Matches 304; Conservative 141; Mismatches 298; Indels 184; Gaps 28;

QY 134 VTE---TQAGEYLLHIQSERANYTVLFTVNVBDTQ-LVYLRPRYFRKMKENQDALLCI-- 186  
 DB 83 IFEKAEATNTGYTC---TNKGGSSSIYVVRDPEKFLIDLPLDLYGKEENDTLVRCLPT 139  
 QY 187 -----SGVPEP-TVEWVLCSSHRESCKEKGPAVVRKEEKVLHFLGTDIRC-- 232  
 DB 140 DPEVTNYSLTGCGEKPLPKDLTFV-----ADPKAGITIRNVKREYHRL-----CLH 185  
 QY 233 CARNALGRE-CTKLFTIDLNOAPOS-----TLPQLFKVGEPLWIRCKALHVNHGFGLT 285  
 DB 186 CSANQKGKMSLKKFKLVRAAIKAPVVSVKTSYLLRGEFEFAVTCLLIKDVSSVDSM 245  
 QY 286 WELEDKALEEGSYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYTSCSSKHPSSALVTI 345  
 DB 246 WIRENSQOTKAQTKKNSHOGDFSYLRQERLTISSARVNDSGVPMCYANNTFGSANVTTT 305  
 QY 346 LE---KGFINA-TSSOEYEDIDPEKFCFSVRKAYPR-IRCTWIFSOASFPCEQRGLED 400  
 DB 306 LEYVDKGFNIFFPMNTTVFVNDGENVDLVVEYEAYPKPHRQWIYNNRT----- 355  
 QY 401 GYSISKFCDHK-----NKPGEYIFYAENDDAQFTKMTNLNIRKK 439  
 DB 356 --STDWDDTPKSENENIRYVVELHLTRLKGTGGTYTHVSNDSVNSVTFNVVNTK 413  
 QY 440 PQVLNANASA--SQASCSGDPPLPSWTWKKCSKSPNCTEEI-PEGVMNKKANKRVFGOW 496  
 DB 414 PEILTHDLVNGMLQCVAGFPPTIDWFCPGTEQRCSPVPGVDVQIQNSSVSPFGKL 473  
 QY 497 VSSSTLNMSEAGKLLVKCCAYNSMGTS-----CETIFLNSGPPFPFI 539  
 DB 474 VVYSTIDDTFKHGTVECRAYNDVGKSSASFNAFKGNSKEQIHAHTLFT-----PLL 527  
 QY 540 QDNISFYATTGLCLPFIWLVILVILCHYKKOFRYESQLOMI-OVTGPLDNEFYVDFRDY 598  
 DB 528 ---IGFVIAAGLMCIFVMIL-----TYKILQKPMTEVQMKVVEELNG---NNVITDPTOL 577  
 QY 599 EYDLKWEFFPRENLEFEGKVLGSGAFGRVMNATAYGISKTGSIQVAVKMLKEKADSCKEA 658  
 DB 578 PYDHKWEFFPRNLSFGKTLGAGAFKGVVEATAYGLKSDAAMTAVKMLKPSAHLTEREA 637  
 QY 659 LMSELKWMTHLGHNDNVNLLGACTLSPGYVLIFFEYCYGDLNLYLSRKRKFKHRTWTEI 718  
 DB 638 LMSELKVLVSLGNHNVNLLGACTIGGPTLVITEYCCYGDLNLFRRKRDSF----- 690  
 QY 719 FKEHNFSSYPTFOAHSSNMPGSRREVOLHPLDQLSGNGNSIHSEDEI-----EYENQ 772  
 DB 691 -----ICSKQEDHAEVALYK-----NLLHKSCESSNDSTNEYMDM 725  
 QY 773 K-----RLAEEEEEDLN-----VLTFEDLLCFAYQVAKGMEFLE 806  
 DB 726 KPGYSYVVPKADKRKSARIGSYIEROVTPAIMEDELALDLEDLLSFYQVAKGMAFLA 785  
 QY 807 FKSCVHRDLAARNVLVTHGKVKVTKCDFGLARDILSDSSYVVRGNARLPVKWMAPELSEG 866  
 DB 786 SKNCIHRDLAARNILLTHGRITKICDFGLARDIKNDSNVYVKGARLPVKWMAPEIFNC 845  
 QY 867 IYTKSDVWSVGIILLWEIFSLGVNPPYPCIPVDANFYKLIQSFKMEQPFYATEGIFYVMO 926  
 DB 846 VYTFESDVWSYGIPLWELFSLGSPYPCMPVDSKPYKMKIEGFRMLSPHAPAEIMDK 905  
 QY 927 SCWAFDSKRKRPSPNLTSLGLCQLAAE 953  
 DB 906 TCWDADPLKRPFTKQIVOLIEKQISES 932  
 RESULT 10  
 KEMS\_RAT  
 ID KEMS\_RAT STANDARD; PRT; 978 AA.  
 AC Q00495;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)

(EC 2.7.1.112) (Fms proto-oncogene) (c-fms).  
 DE CSFIR OR CSFMR OR FMS.  
 GN Rattus norvegicus (Rat).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RC STRAIN=Wistar; TISSUE=Skeletal muscle;  
 RX MEDLINE=93001225; PubMed=1389227;  
 RA Borycki A.G., Guiller M., Leibovitch M.P., Leibovitch S.A.;  
 RT "Molecular cloning of CSF-1 receptor from rat myoblasts. Sequence  
 analysis and regulation during myogenesis.";  
 RL Growth Factors 6:209-218(1992).  
 CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN  
 CC TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: X61479; CAA43706.1; -  
 DR PIR: S16385; S16385.  
 DR HSP: P11362; IFGK.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003598; Ig\_c2.  
 DR InterPro: IPR003600; Ig\_like.  
 DR InterPro: IPR001824; RTKinaselII.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00047; Ig; 4.  
 DR Pfam: PF00069; pkinase; 1.  
 DR ProDom: PD000001; Euk\_pkinase; 2.  
 DR SMART: SM00410; IG\_like; 3.  
 DR SMART: SM00408; IGC2; 1.  
 DR SMART: SM00219; TyrKc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;  
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;  
 KW Immunoglobulin domain; Repeat.  
 FT SIGNAL 1 19  
 FT CHAIN 20 978  
 FT BY SIMILARITY.  
 FT MACROPHAGE COLONY STIMULATING FACTOR I  
 FT RECEPTOR.  
 FT  
 FT DOMAIN 20 511  
 FT TRANSMEM 512 536  
 FT DOMAIN 537 978  
 FT CYTOPLASMIC (POTENTIAL).  
 FT  
 FT DOMAIN 24 104  
 FT IG-LIKE C2-TYPE DOMAIN 1.  
 FT  
 FT DOMAIN 107 197  
 FT IG-LIKE C2-TYPE DOMAIN 2.  
 FT  
 FT DOMAIN 204 298  
 FT IG-LIKE C2-TYPE DOMAIN 3.  
 FT  
 FT DOMAIN 299 397  
 FT IG-LIKE C2-TYPE DOMAIN 4.  
 FT  
 FT DOMAIN 398 503  
 FT IG-LIKE C2-TYPE DOMAIN 5.  
 FT  
 FT DOMAIN 580 914  
 FT PROTEIN KINASE.  
 FT  
 FT NP\_BIND 586 594  
 FT ATP (BY SIMILARITY).  
 FT  
 FT BINDING 614 614  
 FT ATP (BY SIMILARITY).  
 FT  
 FT ACT\_SITE 776 776  
 FT BY SIMILARITY.  
 FT  
 FT DISULFID 42 84  
 FT POTENTIAL.  
 FT  
 FT DISULFID 127 177  
 FT POTENTIAL.  
 FT  
 FT DISULFID 224 278  
 FT POTENTIAL.  
 FT  
 FT DISULFID 417 483  
 FT POTENTIAL.  
 FT  
 FT MOD\_RES 697 697  
 FT PHOSPHORYLATION (AUTO-) (IN VITRO)

FT MOD\_RES 706 706 (BY SIMILARITY).  
 FT PHOSPHORYLATION (AUTO-) (IN VIVO)  
 FT  
 FT MOD\_RES 807 807 (BY SIMILARITY).  
 FT PHOSPHORYLATION (AUTO-) (IN VITRO)  
 FT  
 FT CARBOHYD 45 45 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 73 73 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 302 302 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 335 335 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 389 389 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 410 410 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 449 449 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 478 478 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 491 491 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 978 AA; 109264 MW; 0A68456EF56BC7E3 CRC64;  
 Query Match 23.0%; Score 1210; DB 1; Length 978;  
 Best Local Similarity 33.5%; Pred. No. 1.8e-73;  
 Matches 318; Conservative 147; Mismatches 329; Indels 154; Gaps 31;  
 QY 78 EAATVEAESSITLQVQLATPGDLSCLWVFKHSSLCQPHFDLQNRGIVSMAILNVPTET 137  
 DB 36 ETVTLRCVNGSWEVD-----GPISYWLDPESGS-----TLTRNATFK 77  
 QY 138 QAGEY-LLHIQSERANYTLVTNVNRTQLV-----LRRPYFRKMENQDALL-CIS 187  
 DB 78 NTGYTRCTELEDPMAGSTTI-----HLVYKDPANHSWNLAAQEVTVVEQEAVALPCL- 128  
 QY 188 EGVPEPTVWVLCSSHRESCKEKGPAVVRK-----EKKVLHFLGTDIRCC 233  
 DB 129 --ITDPALK-----DSVSLMRGGRQVLRKTVYFFSAWRGFIIRKAKVL-----DSNTYVC 177  
 QY 234 ARNALGRECTKL-FTIDLN---QAQOSTL-PQLFLVK-GEPLATIRCKAIHNVHGFGLTW 286  
 DB 178 KTMVNGRESTGTGLKVNVRVHPPEPOIKLEPSKLVIRGEAAQIVCSATNAEVGNVIL 237  
 QY 287 ELEDKALE--EGSYFEMSTYSTNRTMIRILAFVSSVGRNDTGYTC--SSSKHPSQSALV 343  
 DB 238 KRGTGKLEIPLNSDFODNYKKVRAL-----SLNAVDFDAGIYSCVASNDVGRTRTATM 291  
 QY 344 T--ILEKGFINATSSQE-EYEIDPYEKFCFSVRKAPRIR-CWIFISOAFCPCBQRLG 399  
 DB 292 NFQVVEAYLNLTSQSLQLQEVSVGDSLTLTVHADAVPSIQHYNNWYLGPEFFE-DORKLE 350  
 QY 400 -----DGYISISKFDH--KNKPEGYIFYAENDDAQFTKMTFLNIRKKPOVLNANASAO 450  
 DB 351 FITORAIYRTFKLFNRVKASEAGQYFLMAQNKAGWNLTFTLTRYPPVSVTMMPVN 410  
 QY 451 AS-----CSSDGYPLPSWTWKCKSDKSPNCTEETEEIPEGVVN---KANKRKVFQGWYSSSTL 502  
 DB 411 GSDVLFCDVSGYPOPSVTWMECRGHTDRCDQAQALQWWDTHPEVLSOKPGDKVLIQSQL 470  
 QY 503 NMSEAGKGLVKKCCAYNSMGTSCETIFLNSPGPPFFIODNISFYATIGLC---LPFIIVL 559  
 DB 471 PIGTLKHNMTYFCKTHNSVGNSSQYFRAVSLGQSKQLPDESILFTPVVACSVMSLLVLL 530  
 QY 560 IVLICHYKKOFRYESQLQMIQ-VTGPLDNEYFYDFRDYEDLKWEPFRENLEFGKVLG 618  
 DB 531 LLLLYKYKQPKYQVRWKIERYEG---NSYTFIDPTQLPYNEKWEPPRNALQFGKTLG 587  
 QY 619 SGAFGRVNNATAYGISKTGVSIQVAVKMLKEKADSCERKALMSLKMTHLGHHDHNVNL 678  
 DB 588 AGAFGKVVEATAFGLGKEDAVLKVAVRMLKSTAHADKALMSLKMTHLGHHDHNVNL 647  
 QY 679 LGACTLSGPVYLIFEYCCYGDLLNVLRSKRE-----KFHRTWTET 718  
 DB 648 LGACTHGGPVLVITEYCCYGDLLNVLRSKRE-----KFHRTWTET 707  
 QY 719 FKEHNFSSYPTFOAHNSNMPGSRREVQLPPLDQLSGFNNGNSIHSEDEIYEYENQKLAE 778  
 DB 708 RRDGSGFS-----SQGVDTYVEMRPVSTSSDSFFKQD-----L 740  
 QY 779 EEDLNVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGVKVKICDFGLARD 838

```
Db 741 DREPSRPLEWDLHFFSSQAOCMAFLASKNCIHRDVAARNVLLTSGHVAKIGDFGLARD 800
Qy 839 ILSDSYVVGNGARLPVKWAPESLFEGETYIKSDVWSYGILLWEIFSLGVNYPGPVD 898
Db 801 IMDSNYVVGNGARLPVKWAPESILYCYTVQSDVWSYGILLWEIFSLGLNYPGILVN 860
Qy 899 ANFYKLIQSGKMEQPFYATEGYFYVMQSWAFDSDKRPSFPLNTSFL 946
Db 861 NKFYKLVKDGQYMAQPFAPKNIYSIMQSCWDLPTRRPTFOQICPLL 908

RESULT 11
KIT_CANFA
ID KIT_CANFA STANDARD; PRT; 975 AA.
AC 097799;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (proto-oncogene tyrosine-protein kinase kit) (c-kit).
GN KIT.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99142897; PubMed=9989791;
RA Ma Y., Longley B.J., Wang X., Blount J.L., Langley K., Caughey G.H.;
RT "Clustering of activating mutations in c-KIT's juxtamembrane coding
RT region in canine mast cell neoplasms.";
RL J. Invest. Dermatol. 112:165-170(1999).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY..
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF044249; AAD002327.1; .
CC HSP; P11362; 1FGK.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR003599; Ig.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003600; Ig_Like.
CC InterPro: IPR001824; RTKinaseIII.
CC InterPro: IPR001245; Tyr_pkinase.
CC Pfam: PF00047; Ig; 2.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Euk_pkinase; 2.
CC SMART; SM00409; IG; 2.
CC SMART; SM00410; IG_Like; 1.
CC SMART; SM00219; TYRK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
CC Transferase; Glycoprotein; Phosphorylation; ATP-Binding; Signal;
```

```
Immunoglobulin domain.
KW SIGNAL 1 24
FT CHAIN 25 975
FT DOMAIN 25 519
FT TRANSMEM 520 542
FT DOMAIN 543 975
FT DOMAIN 588 936
FT NP_BIND 622 622
FT BINDING 622 622
FT ACT_SITE 791 791
FT MOD_RES 822 822
FT CARBOHYD 96 96
FT CARBOHYD 132 132
FT CARBOHYD 147 147
FT CARBOHYD 286 286
FT CARBOHYD 296 296
FT CARBOHYD 303 303
FT CARBOHYD 355 355
FT CARBOHYD 370 370
FT CARBOHYD 403 403
FT CARBOHYD 466 466
FT CARBOHYD 489 489
FT SEQUENCE 975 AA; 109335 MW; 8F570BDB9F05B1CB CRC64;

Query Match 22.9%; Score 1207; DB 1; Length 975;
Best Local Similarity 31.9%; Pred. No. 2.8e-73;
Matches 330; Conservative 165; Mismatches 375; Indels 166; Gaps 34;

Qy 31 LPVIKCVLISHENNGSSAGKPSYRMVGRSGPEDLOCTPRROSEGTVVEAAATVEAESGSI 90
Db 11 LCVLLLLLLGVRGTSS--QPS-----VSPGEPSPSLPSIHPAKSELIVSVGD- 54
Qy 91 TLQVOLATPGDLSCLVWFKHSSLCGCPHFQDLQNRGIVSMALNVETQOAGYLLHIOQSR 150
Db 55 ELRLSCTDPGEVK--WTFE--TLG-QLNENTHNEWITEKA-----EAGHTGNYCTNR 102
Qy 151 ANYTVLFTVNVDR--TQLYVLRPRYFRKMENODALLCISEGVPEPTVEVWLSCSHRESCKE 209
Db 103 DGLSRSYVTFVDPKALFLVDLPYLGKRGNDTLVRC-----PLTDPVETNYSLRGC-- 153
Qy 210 EGPVAVRKEEKVLHELFCTDIR-----C--CARNALGRE-CTKLFTIDLNOAPOS- 256
Db 154 EGPPLKDLTFVADPKAGITIRNVKREYHRLCLHCSADOKGRTVLSKFTLVKRAIRAV 213
Qy 257 -----TLPLQLFKVGEPLWIRCKAIHVNHGFGLTWELEDKALEEGSYFEMSTYSTNRTMI 311
Db 214 PVSVSKTSSLLKGEAFVSMCFIKDYSSFVDSMWIKENSQOTNAQTOSNSWHGDFNFE 273
Qy 312 RILAFVSSVGRNDTGYTTCSSKHPQSALVTILE---KGFINA-TSQEEYEIDPYEK 367
Db 274 ROEKLIITSSARVNDSGVFMVCYANNTTFTLEVVDKGFINFPMHSTTIFYNDGON 333
Qy 368 FCFVSRKAYPR-IRCTWIFSQASF--PCEQRGLDGYISIKFCDH-----KNKPGCY 417
Db 334 VDLIVEYAYPKPHQOQIYNNRTFTDKWEDYPKSDNESNIRYVSELHLTRLKGEGTY 393
Qy 418 IFYAENDDAQFTKMFNTLNIIRKPKQVLANASQA--SCSSDGYPLSPWTWKCKSDKSPNC 475
Db 394 TFQVNSDVNSSTVFNVYVNTKPEILLTHESLTNGMLQCVAGFPPEAVGVWFCPGAQRK 453
Qy 476 TEEI-PEGVWNKANKRVFGOWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCETIFLNSPG 534
Db 454 SVPIGPMQVQMSLSLSPSGKLVQSSIDYSAFKINGTVECRAYNNVGRS--SAFEN--- 508
Qy 535 PFPIQDNIIFYATIGLCLPFIV-----VLVLICHYKKKOFRYVESOLOMI-QVTGPLD 587
Db 509 -FAPKEQIHPHTLFTPLLIGEVIAAGMMCIIVNLTLYLQKPMYEVQNVKEELNG--- 564
Qy 588 NEYFVVDPRDYEDLKWEPFRENLEFGKVLGSGAFGRVNNATAYGISGTGSIQVAVKML 647
Db 565 NNYVIIDTQPLPDHKWEPFNRNLSFGTKLAGAGFKVVEATAYGLIKSDAAMTAVKML 624
Qy 648 KEKADSCKEALMSSELKMMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLSK 707
```

```
Db 625 KPSAHLTEREALMSKLVSYLGNHNNIYNLLGACTVGGPTLVITEYCCYGDLLNPLRRK 684
Qy 708 REKFHTWTEIFKEHNFSSYPFOAHSSNMSPGSRVQLHPPLDQLSGNGSIHSEDEI 767
Db 685 RDSFICSQOE---DHG-----EVALYK-----NLHSHKSS 712
Qy 768 -----EYENOK-----RLAEEDLNL-----VLTFEDLLCFA 795
Db 713 CSDSTNEYMDMPGVSYVVPTRADKRRSRIGSYIERDVTPALMEDELALDELDSFS 772
Qy 796 YQVAKGMFLEPKSCVHRDLARNVLVTHGKVVKICDFGLARDIILSDSYVVRGNARLPV 855
Db 773 YQVAKGMFLASKNGICHRDLARNILLTHGRITKICDFGLARDIKNDSYVVRGNARLPV 832
Qy 856 KMWAPESLFEGYITTKSDVMSYGLLWELFSLGVNPGIPVDANFYKLIOGFKMEQPF 915
Db 833 KMWAPESINFCVITTESDVMWSYGIFLWELFSLGVNPGIPVDANFYKLIOGFKMEQPF 892
Qy 916 VATEGIYFVMSCAFDSRKRPSFNLTLFGLQGLAEAEACIRTSIHLPKQAAPQORG 975
Db 893 HAPAEYDIMTKCWDADPLKRTFTKQIVOLIEKQISDSIN-----HIYSNLAN----- 940
Qy 976 LRAQSPQROVKTHRR 991
Db 941 -CSPNPERPVVDHSVR 955

RESULT 12
KIT_HUMAN
ID KIT_HUMAN STANDARD; PRT: 976 AA.
AC P10721;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit) (CD117 antigen).
GN KIT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Term placenta, and Fetal brain;
RX MEDLINE=88111521; PubMed=2448137;
RA Yarden Y., Kuang W.-J., Yang-Feng T., Coussens L., Munemitsu S.,
RA Dull T.J., Chen E., Schlessinger J., Francke U., Ullrich A.;
RT "Human proto-oncogene c-kit: a new cell surface receptor tyrosine
RT kinase for an unidentified ligand.";
RL EMBO J. 6:3341-3351(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93064697; PubMed=1279499;
RA Giebel L.B., Strunk K.M., Holmes S.A., Spritz R.A.;
RT "Organization and nucleotide sequence of the human KIT
RT cell growth factor receptor) proto-oncogene.";
RL oncogene 7:2207-2217(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97179223; PubMed=9027509;
RA Andre C., Hampe A., Lachaume P., Martin E., Wang X.P., Manus V.,
RA Hu W.X., Galibert F.;
RT "Sequence analysis of two genomic regions containing the KIT and the
RT FMS receptor tyrosine kinase genes.";
RL Genomics 39:216-226(1997).
RN [4]
RP VARIANT LYS-583.
RX MEDLINE=92291284; PubMed=1376329;
RA Fleischman R.A.;
RT "Human piebald trait resulting from a dominant negative mutant allele
RT of the c-kit membrane receptor gene.";
RL J. Clin. Invest. 89:1713-1717(1992).

[5]
RN RP VARIANT LEU-584.
RX MEDLINE=92133600; PubMed=1370874;
RA Spritz R.A., Giebel L.B., Holmes S.A.;
RT "Dominant negative and loss of function mutations of the c-kit
RT (mast/stem cell growth factor receptor) proto-oncogene in human
RT piebaldism.";
RL Am. J. Hum. Genet. 50:261-269(1992).
RN [6]
RN RP VARIANT ARG-664.
RX MEDLINE=92020918; PubMed=1717985;
RA Giebel L.B., Spritz R.A.;
RT "Mutation of the KIT (mast/stem cell growth factor receptor)
RT protooncogene in human piebaldism.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:8696-8699(1991).
RN [7]
RN RP VARIANT VAL-816.
RX MEDLINE=94013473; PubMed=7691885;
RA Furitsu T., Tsujimura T., Tono T., Ikeda H., Kitayama H.,
RA Koshimizu U., Sugahara H., Butterfield J.H., Ashman L.K.,
RA Kanayama Y., Matsuzawa Y., Kitamura Y., Kanakura Y.;
RT "Identification of mutations in the coding sequence of the proto-
RT oncogene c-kit in a human mast cell leukemia cell line causing
RT ligand-independent activation of c-kit product.";
RL J. Clin. Invest. 92:1736-1744(1993).
RN [8]
RN RP VARIANTS PIEBALDISM GLY-791 AND VAL-812.
RX MEDLINE=93322624; PubMed=7687267;
RA Spritz R.A., Holmes S.A., Itin P., Kuester W.;
RT "Novel mutations of the KIT (mast/stem cell growth factor receptor)
RT proto-oncogene in human piebaldism.";
RL J. Invest. Dermatol. 101:22-25(1993).
RN [9]
RN RP VARIANT PIEBALDISM 893-GLU--PRO-896 DEL.
RX MEDLINE=96287384; PubMed=8680409;
RA Riva P., Milani N., Gandolfi P., Larizza L.;
RT "A 12-bp deletion (7818del12) in the c-kit protooncogene in a large
RT Italian kindred with piebaldism.";
RL Hum. Mutat. 6:343-345(1995).
RN [10]
RN RP VARIANT GIST VAL-559 DEL.
RX MEDLINE=98361155; PubMed=9697690;
RA Nishida T., Hirota S., Taniguchi M., Hashimoto K., Isozaki K.,
RA Nakamura H., Kanakura Y., Tanaka T., Takabayashi A., Matsuda H.,
RA Kitamura Y.;
RT "Familial gastrointestinal stromal tumours with germline mutation of
RT the KIT gene.";
RL Nat. Genet. 19:323-324(1998).
RN [11]
RN RP FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
RN GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
RN BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
RN ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
RN 3-KINASE (PI3K).
RN [12]
RN RP CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
RN tyrosine phosphate.
RN [13]
RN RP SUBCELLULAR LOCATION: Type I membrane protein.
RN [14]
RN RP DISEASE: DEFECTS IN KIT ARE THE CAUSE OF A DEVELOPMENTAL
RN ABNORMALITY KNOWN AS PIEBALDISM. AN AUTOSOMAL DOMINANT GENETIC
RN DISORDER OF PIGMENTATION CHARACTERIZED BY CONGENITAL PATCHES OF
RN WHITE SKIN AND HAIR THAT LACK MELANOCYTES.
RN [15]
RN RP DISEASE: DEFECTS IN KIT ARE THE CAUSE OF A GASTROINTESTINAL
RN STROMAL TUMOR (GIST).
RN [16]
RN RP SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
RN PROTEIN KINASES.
RN [17]
RN RP SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
RN [18]
RN RP DATABASE: NAME=PROW; NOTE=CD guide CD117 entry;
RN WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd117.htm".
RN [19]
RN RP DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
RN WWW="http://www.infobiogen.fr/services/chronocancer/Genes/KIT10127.html".
RN [20]
RN RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL Outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
```



Db 779 GNAFLASKNCIHRDLAARNILLTHGRITICDFGLARDIKNDNSYVYKGNARLPVKWAP 838  
 QY 861 ESLFEGIVTIKSDVWSYGLLEWELSLGVPYGPVDFANFYKLLIOSGKMQPFFATEG 920  
 Db 839 ESIFNCVTFESDWSYGLFLWELSLGSPYGMVDFSKFYKMIKEGFRMLSPHEAPAE 898  
 QY 921 IYFVMSQWAFDSRRKSPNPNTSLFGLQALAE 953  
 Db 899 MYDIKTCWDADPLKPTKQIVOLIEKOISE 931

RESULT 13  
 KIT\_CHICK STANDARD; PRT; 960 AA.  
 AC 008156;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, last sequence update)  
 DT 16-OCT-2001 (Rel. 40, last annotation update)  
 DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)  
 DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).  
 GN KIT.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=white leghorn; TISSUE=Brain;  
 RX MEDLINE=93292995; PubMed=7685729;  
 RA Sasaki E., Okamura H., Chikamune T., Kanai Y., Watanabe M.,  
 RA Naito M., Sakurai M.;  
 RT Cloning and expression of the chicken c-kit proto-oncogene.";  
 RL Gene 128:257-261(1993).  
 CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL  
 CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.  
 CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND  
 CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL  
 CC 3-KINASE (PI3K).  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: HIGH IN THE BRAIN AND TESTES AND ALSO  
 CC PRESENT IN THE BURSA OF FABRICUS, HEART, KIDNEY, LUNG, SPLEEN  
 CC THYMUS AND OVARY.  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; D13225; BAA02506.1;  
 CC PIR; JN0677; JN0677.  
 CC HSSP; P11362; IFGK.  
 CC InterPro; IPR000719; Euk\_pkinase.  
 CC InterPro; IPR003598; Ig\_c2.  
 CC InterPro; IPR01824; RTKinaseIII.  
 CC InterPro; IPR01245; Tyr\_pkinase.  
 CC Pfam; PF00047; Ig\_3.  
 CC Pfam; PF00069; pkinase; 1.  
 CC ProDom; PD000001; Euk\_pkinase; 2.  
 CC SMART; SM00408; Igc2; 1.  
 CC SMART; SM00219; TyrKc; 1.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.

DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;  
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;  
 KW Immunoglobulin domain.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CHAIN 25 960 MAST/STEM CELL GROWTH FACTOR RECEPTOR.  
 FT DOMAIN 25 505 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 506 530 POTENTIAL.  
 FT DOMAIN 531 960 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 575 913 PROTEIN KINASE.  
 FT NP\_BIND 581 589 ATP (BY SIMILARITY).  
 FT BINDING 609 609 ATP (BY SIMILARITY).  
 FT ACT\_SITE 777 777 BY SIMILARITY.  
 FT MOD\_RES 808 808 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 76 76 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 149 149 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 269 269 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 286 286 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 306 306 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 318 318 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 338 338 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 343 343 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 356 356 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 453 453 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 469 469 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 960 AA; 107311 MW; 0E93850527AB68F6 CRC64;  
 Query Match 22.4%; Score 1181; DB 1; Length 960;  
 Best Local Similarity 31.8%; Pred. No. 1.5e-71;  
 Matches 310; Conservative 159; Mismatches 356; Indels 150; Gaps 30;  
 QY 45 GSSAGKPPSYRVGRGSPEDLOCTPRQSEGTVEATVEAECSITLQVLATPGDLSC 104  
 Db 24 GSPVHEESSLVNKGELRLKN-----EGPVTWNFQNSDPSAKTR 65  
 QY 105 L-----WPKHSLGCPHDFDLQNRGIVSMAILNVNTEQAGEYLLHIQSERANYTVLFTV 159  
 Db 66 ISNEKEWHYKNAITRIDIGRYECKSG-----SIVN-----SFYV 99  
 QY 160 NVBTDQLYVLRPYFRKMENQDALLCISEGVPETVEWILCSSHRESCKEGBGPAVKEE 219  
 Db 100 FVKDPNVFLVDSLIYKEDSDILLVCLTDPD-VLNFTLRKCKDGKPLKPMTEIPNPK 158  
 QY 220 ----KVLHELFGTDIRCCAR-NALGRECTKLFTIDLNQAP-QSTLPOL-----FLKVG 266  
 Db 159 GIILKNVQRSGKCYCQCLAKHNGVEKISEHIF---LNVRPVHKALPVITLTKSYELLKEG 215  
 QY 267 EPLWIRCKATHVNHGFLTWELEDKALEGSYFEMSTYSTNRMTIRILLAFVSSVGRNDT 326  
 Db 216 EEEFVTCITDVDSYKASHYSKSAIVTSKRNLDYGYERK----LTNIRSVGVNDS 271  
 QY 327 GYITCSSKHP--SQSALVTI--LEKGFINATSSQE-EYIDPYEKFCFSVRKAYPRIR 381  
 Db 272 GEFTC-QAENPFCKTNATVTKALAKEGVRLEFATMTNTIDINAGONGNLTVEYEAYPK-- 328  
 QY 382 CTWIFSQAFPCQORLEGDYSISKCDHNK-----PGEYIF 419  
 Db 329 -----PKEEVVMYMNNTLQNSDHYVFKTVKVNNSYTSSELHLTRLKGTGGIYTF 378  
 QY 420 YAENDDAQFTKMTLNLIRKKPOVLANASASO--ASCSSDGYPLPSWTWKCKSDKSPNCTE 477  
 Db 379 FVNSDASSVTENVYKTKPEILTLDMLGNILQCVATGFPAPTIIWYFCPGTEQRCLD 438  
 QY 478 E----IPEGVWNKKANRKY--FGQWVSSSTLNMSEAGKGLLVKCCAYNSMGTSCTEIFLNS 532  
 Db 439 SPTISPMDVKVSYTNSVSPSFERILVESTVNAS-MFKSTGTICCEASNSGDK-SSVFNENF 496  
 QY 533 PGFPFIQDNISFYA-----TIGLCPLFIVLVILVILCHYKQFYRESOLQMI-QVTG 584  
 Db 497 A-----IKEQIRHTLTFTPLLIAGVAAAGLMCIIVMLVYIYLOKPKYEVQWVVEEING 551  
 QY 585 PLDNEYFYVDYFRDEYDLKWEFFRENLEFKVLGSGAGFRVMNATAYGISKTGSIQAV 644



Db 552 ---NNVYVDPQLPYDKWEPNRLSFGKTLGAGAGKVVVEATAYGLFKSDAAMTAV 608  
Qy 645 KMLKEKADSCKEALMSKLMWTHLGHNDINVLNGLAGTSLGPPVLYEYCYGDLNVL 704  
Db 609 KMLKPSAHLTEREALMSKLSYLGNIHINVLNGLAGTIGGTLVITEYCYGDLNVL 668  
Qy 705 RSKREKF---HRTWTE--IFKEHNFSSYPTFOA-HSNSSMPGSRREVQLHPPLDQSGFN 757  
Db 669 RKRSDFICPHEEBAEAAVYENLLHQAEPTADAVNEYMDKPGVSYAVPPKADKRPVK 728  
Qy 758 GNSIHSEDEIEYENOKRLAEEDENLVTFFDLCLCFAYQVAKGMEFLFKSCVHRLAA 817  
Db 729 SGS-----YTDQDVTLSMLEDDDELALDVEDLLSFSYQVAKGMSFLASKNCIHRDLAA 780  
Qy 818 RNVLTGKVKVVICDFGLARDILSDSSVYVGNARLPVKWMAPELFCGIYTIKSDVWSY 877  
Db 781 RNILTHGRITKICDFGLARDIRNSYVYVGNARLPVKWMAPELFCNFVYTFESDVWSY 840  
Qy 878 GILLWEIFSLGPNVPGIPVDANFYKLIQSGFKMQPFFVATEGIVFMQSWAFDSRRKP 937  
Db 841 GILLWELSLGSPPGMPVDSKFKYMIKEGYRMFSPESCPPEMYDIMKSCWDADPLQRP 900  
Qy 938 SPPNLTSLFLGQALAE 952  
Db 901 TFKQIVOLTEQQLSD 915  
RESULT 14  
KIT\_FELCA  
ID KIT\_FELCA STANDARD; PRT; 978 AA.  
AC Q28889;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)  
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).  
KIT  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
OX NCBI\_TaxID=9685;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Head;  
RX MEDLINE=95140426; PubMed=7530827;  
RA Herbst R., Muneimitsu S., Ullrich A.;  
RT "Oncogenic activation of v-kit involves deletion of a putative  
RT tyrosine-substrate interaction site.";  
RL Oncogene 10:369-379(1995).  
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL  
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.  
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND  
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL  
CC 3-KINASE (PI3K).  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
CC tyrosine phosphate.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
CC PROTEIN KINASES.  
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; S76596; AAB33207.1;  
CC HSSP; P11362; IFGK.  
DR InterPro; IPR000719; Euk\_pkinase.

DR InterPro; IPR003599; Ig.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003600; Ig\_like.  
DR InterPro; IPR001824; RTKinaseIII.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00047; Ig; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 2.  
DR SMART; SM00409; IG; 2.  
DR SMART; SM00410; IG\_like; 1.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;  
KW Transferrase; Glycoprotein; Phosphorylation; ATP-binding; Signal;  
KW Immunoglobulin domain.  
FT SIGNAL; 1 22  
FT CHAIN; 23 978  
FT DOMAIN; 23 521  
FT TRANSMEM; 522 544  
FT DOMAIN; 545 978  
FT DOMAIN; 590 938  
FT NP\_BIND; 596 604  
FT BINDING; 624 624  
FT ACT\_SITE; 793 793  
FT MOD\_RES; 824 824  
FT CARBOHYD; 94 94  
FT CARBOHYD; 145 145  
FT CARBOHYD; 284 284  
FT CARBOHYD; 294 294  
FT CARBOHYD; 301 301  
FT CARBOHYD; 321 321  
FT CARBOHYD; 353 353  
FT CARBOHYD; 368 368  
FT CARBOHYD; 401 401  
FT CARBOHYD; 464 464  
FT CARBOHYD; 487 487  
SQ SEQUENCE 978 AA; 109449 MW; 6D45472E07440E6B CRC64;  
Query Match 22.4%; Score 1178.5; DB 1; Length 978;  
Best Local Similarity 31.8%; Pred. No. 2.3e-71;  
Matches 313; Conservative 162; Mismatches 373; Indels 143; Gaps 30;  
Qy 36 CVLISHENNGSSAGKPSYRMVRGSPEDLOCTPRRQSEGVYEAATVEAEGSGITLOVQ 95  
Db 12 CVLLLLRVTGSSQPSA-----SPGWSLPSIH-PATSELIYVSGADEIRLL 57  
Qy 96 LATPGDLSCLVFKHSSLGCGQPHFDLQNRGIVSMAILNVETQAGEYLLHIQSERANTV 155  
Db 58 CTDPGFVK--WTPE--TLG-QSSEITHNEWITEKA---EATNTGNYTC---TNGGGGLSS 105  
Qy 156 LFTVNVDR-TQLVYLRPRPFYKMNQDALLCI-----SEGVPPEP----- 193  
Db 106 SIYFVDRPAKFLVLDLPLGYKEDHDLVRCPLTDPVNTVSLRGCEGCKPLPKLFTFVD 165  
Qy 194 -----TVEWVLCSSHR--ESCKEKGPAVVRKEVHLFETDIRCCARNALGRECTKLF 246  
Db 166 PKAGITIRNVKREYHRLCHCSAD-----RAGKSVLSKKFTLVRAAIR-----AVP 212  
Qy 247 TIDLNOAQPOSTLPOLFLVKGLPELWIRCKAIHVNHGFGTLWELEKALEEGSYFENSTYST 306  
Db 213 VVSYSKASH-----LLREGEFEFSVMCLIKDYSSSDVSDSMWIKENSPTNAOPQSNWHQ 266  
Qy 307 NRTMIRILLAEVSSVGRNDTGYTCSSSKHPSQSALVILE---KGFINA-TSSOEEVEI 362  
Db 267 DFNVRQERLTISSARVNDSGVFMCIYANNTFGSANVTTLEVYAKGFIIPFMMNTTIFV 326  
Qy 363 DPYEKFCFSYRFPAYPR-IRCTWIFSOASFCEQRLGDEGYSISKFCDHK----- 411  
Db 327 NNGENVLDVIEYAYPKPEHORWVYM-----NRTLTDKWDYDKPSDNESNIRYVSELH 379



QY 212 P-----AVVRKEKVLHELFGTDIROCARNALRECTKLFETIDLNOAQSTLPQLFLKVG 267  
Db 185 PYICATVR-----GRTFKTSFNFYALKATSELNLEMDIR-----QTVYKAGE 228  
QY 268 PLWIRCKAIHVNHGFLWLEDEKALEGSEYFEMSTYSTRNRMIR-----ILLAFVSV 321  
Db 229 TIVVTC-AVFNEVVVDLQWTPYGEVRNGI-----TMLEEIKLPSIKLVYTLTV 276  
QY 322 GR--NDGCTYTCSSSKHPSQSALVTIL--EKGFINATSQBEYE-IDPYEKFCFSV 372  
Db 277 PKATVKDSDGYECARQATKEVKEMKTVTIVSHERGVOIRPTFGHLETVNLNHQVREFV 336  
QY 373 RFKAYPRIRCTWIFSOASP-----PCEQRGLEGY-SISKFCDHKNK-PGEYIFYAE 422  
Db 337 EQVAYPTPRISWLKDNLTILNLTITTTDVORSQETRYOSKLLIRAKEEDSGHYTIIVQ 396  
QY 423 NDDAFTKMTFTINIRKKPOVL-----ANASQASCSSDGYPLPSWTW-----KKCS 469  
Db 397 NDDMKSVTFELSTLVPASILELVDHGGGGGTVRCTAEGTPLFNIEWMICKDIKKCN 456  
QY 470 DK-----SPNCTEETPEGVWKKANRKYFGQWVSSSTLNMSEAGKGLLVKCCAYNSMGT 523  
Db 457 NDTSTWTLASVNSNIITE--FHQRGRSTVEGR-----VSFAKVEETIAVRCLAKNDLGI 508  
QY 524 SCETIFLNSPGPPFTQDNISFYATIGLCLPFIVLVILVILCHYKKOFRYESOLOMIQVT 583  
Db 509 GNRELKLVAPS-----LRSELTVAADVLLVIVIVSVLIVLVIVWOKPRYEIRWRVIESI 564  
QY 584 GPLDNEYFYVDRDYEDLKWFFPRENLNLEFGKVLGSGAFGRVNNATAYGISKTGVSIOVA 643  
Db 565 SPDGHEYIYVDPMLPYDSRWEFFPRDGLVLRILGSGAFGKVVVEGTAYGLSRSQPVKVA 624  
QY 644 VKMLREKADSCKEALMSLKMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNY 703  
Db 625 VKMLKPTARSEKQALMSLKMTHLGHPLNIVNLLGACTKSGPIYIITEYCFYGDLVNY 684  
QY 704 LRSKREKF-----HRTWTEIFKEHNFSSYPTFOAHSNSSMP- 739  
Db 685 LHKNRDSFMSRHPKPKDLDFGLNPADESTRSYVILSFENNGDYVDMKQADTTQYVPM 744  
QY 740 -GSREVQLHPDLQSGFNGNSIHSEDEIEYENOKRLAEEDLNLVTFEDLLCFAYQV 798  
Db 745 LERKEVSKYSDIQRSYDRPASVYKKKMLDSEAKNLLSDDDSEGLTLL---DLSFTYQV 801  
QY 799 AKGMEFLEFKSCVHRDLAARNVLVTHGKVKVTCDFGLARDILSDSSYVVRGNARLPVKWM 858  
Db 802 ARGMEFLASKNCVHRDLAARNVLLAQGKIVKICDFGLARDIMHDSNYSVSKGSTFLPVKWM 861  
QY 859 APESLFEGIYTTIKSDVMSYGILWEIFSLGVNPPYGPVDPANFYKLIQSGFKMEQPPYAT 918  
Db 862 APESIFDNLYTTLSDVMSYGVLWEIFSLGGTYPGCMWVDSTFYNKIKSGYRMAKPDHAT 921  
QY 919 EGIYFVMSQAFDSKRKPSFPNLTSL 946  
Db 922 SEVYEIMVQWNSPEKEKPSFYHLSEIV 949

Search completed: May 27, 2003, 14:32:10  
Job time : 42.9909 secs

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:26:24 ; Search time 70.9642 Seconds

(without alignments)  
2880.307 Million cell updates/sec

Title: US-09-919-408-2

Perfect score: 5264

Sequence: 1 MRALAQRSDRRLLLVLSV.....RGGLRAQSPQVQKIHRS 992

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1268	24.1	976	13 Q9W755	Q9W755 brachydanio
2	1234.5	23.5	974	11 Q63702	Q63702 rattus ratt
3	1232.5	23.4	978	11 Q63116	Q63116 rattus norv
4	1220	23.2	954	13 Q91909	Q91909 xenopus lae
5	1215.5	23.1	974	13 Q985U3	Q985U3 danio dangi
6	1215.5	23.1	977	13 Q985U1	Q985U1 danio nigro
7	1214	23.1	977	13 Q918N6	Q918N6 brachydanio
8	1209.5	23.0	979	6 Q8WN23	Q8WN23 canis fami
9	1208	22.9	964	6 Q97744	Q97744 sus scrofa
10	1207	22.9	964	6 Q9TQ01	Q9TQ01 sus scrofa
11	1207	22.9	964	6 Q9TQ00	Q9TQ00 sus scrofa
12	1205	22.9	978	6 Q9XS93	Q9XS93 canis fami
13	1204.5	22.9	977	13 Q985U2	Q985U2 danio kerri
14	1201	22.8	948	6 Q9T7D7	Q9T7D7 trichosurus
15	1198.5	22.8	977	13 Q985U4	Q985U4 danio albol
16	1194	22.7	972	4 Q99662	Q99662 homo sapien

17 1188.5 22.6 975 13 P79750  
18 1173 22.3 945 6 O77589  
19 1170.5 22.2 724 6 Q9MYN0  
20 1136 21.6 1059 13 Q9DE49  
21 1126.5 21.4 1087 13 Q9PUF6  
22 1110 21.1 1097 11 Q8R406  
23 1073 20.4 1019 13 Q8UV88  
24 1060 20.1 1048 13 P79749  
25 1013.5 19.3 986 13 Q8UV89  
26 997.5 18.9 923 6 O77745  
27 980 18.6 1379 13 P79701  
28 974 18.5 1327 13 Q8OHL3  
29 961 18.3 1363 11 Q91ZT1  
30 957 18.2 1301 13 Q8U0W9  
31 946 18.0 1173 13 Q9PTL0  
32 939 17.8 563 11 Q925F7  
33 937.5 17.8 1345 11 Q8VCD0  
34 901.5 17.1 323 11 Q9EO22  
35 877.5 16.7 323 11 Q9EQ24  
36 833 15.8 345 13 Q9PVU7  
37 820.5 15.6 350 13 Q91416  
38 798 15.2 796 13 Q91287  
39 793 15.1 766 4 Q96KM2  
40 793 15.1 785 4 Q96KM1  
41 793 15.1 819 4 Q96KM0  
42 791 15.0 824 13 Q90749  
43 790.5 15.0 806 13 Q90Z00  
44 790.5 15.0 820 4 Q96KL9  
45 788.5 15.0 922 13 Q90413

## ALIGNMENTS

RESULT 1

Q9W755 PRELIMINARY; PRT; 976 AA.  
ID Q9W755  
AC Q9W755;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Kit receptor tyrosine kinase.  
GN KIT.  
OS Brachydanio rerio (Zebrafish) (Zebra danio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99396707; PubMed=10393121;  
RA Parichy D.M., Rawls J.F., Pratt S.J., Whitfield T.T., Johnson S.L.;  
RT "Zebrafish sparse corresponds to an orthologue of c-kit and is  
RT required for the morphogenesis of a subpopulation of melanocytes, but  
RT is not essential for hematopoiesis or primordial germ cell  
RT development."  
RL Development 126:3425-3436(1999).  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
CC PROTEIN KINASES.  
DR EMBL: AF153446; AAD41890.1; -.  
DR HSSP: P11362; IFCK.  
DR ZFIN: ZDB-GENE-980526-464; kit.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR003600; Ig\_Like.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR001824; RTKaseIII.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR Pfam: PF00047; Ig\_3.  
DR Pfam: PF00069; Pkinase; 1.  
DR ProDom: PD000001; Euk\_pkinase; 2.  
DR SMART: SM00408; IGC2; 1.

P79750 fugu rubrip  
O77589 equus caball  
Q9MYN0 bos taurus  
Q9DE49 brachydanio  
Q9PUF6 gallus gall  
Q8R406 rattus norv  
Q8UV88 fugu rubrip  
P79749 fugu rubrip  
Q8UV89 fugu rubrip  
O97745 sus scrofa  
P79701 coturnix co  
Q8OHL3 gallus gall  
Q91ZT1 rattus norv  
Q8U0W9 brachydanio  
Q9PTL0 brachydanio  
Q925F7 rattus norv  
Q8VCD0 mus musculu  
Q9EQ22 rattus norv  
Q9EQ24 rattus norv  
Q9PVU7 lethenteron  
Q91416 xenopus lae  
Q91287 pleurodeles  
Q96KM2 homo sapien  
Q96KM1 homo sapien  
Q96KM0 homo sapien  
Q90749 gallus gall  
Q90Z00 brachydanio  
Q96KL9 homo sapien  
Q90413 brachydanio

DR SMART; SM00410; IG-like; 1.  
DR SMART; SM00219; TYKIC; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
KW Phosphorylation; Receptor; Transferase; Transmembrane;  
KW Tyrosine-protein kinase.  
SQ SEQUENCE 976 AA; 109278 MW; A80AA01658C1A2DA CRC64;  
  
Query Match 24.1%; Score 1268; DB 13; Length 976;  
Best Local Similarity 36.6%; Pred. No. 1.5e-102;  
Matches 318; Conservative 138; Mismatches 302; Indels 110; Gaps 29;  
  
Qy 177 MENODALLICSEGPPEPVENWVLCSSHRESCKEEGPAV--VRKEKVLHLEFGTDIRCCA 234  
Db 138 MKNLNQK--DGQPLPN-----SLRYSASLETGVSQVKRKEFGYVCVGTLDAAATV 189  
Qy 235 RNALGRECTKLTIDNOAPOSTLP-----QLFLKVGPELWIRCKAIHVNHGFLTWE 287  
Db 190 KS--GR-----YQLTVRLVPPAPPTITLQGPORVLLTQGEKLSLCSSTSNVNSDIATVKKW 242  
Qy 288 LE---DKALEEGSFEMSTYTNRTMIRILLAFVSSVGRNDGYTCSSSKHPSOSA--- 341  
Db 243 APNGVNPVSHQNSHL---LPEPIHVRTAILSLSSVTMQDAGNYSCEAINEKGTAKPV 298  
Qy 342 LVITILEKEFINATS--SQEYEDIPYEKFCFVRKAYPRIC--TWIFSOASPPCEQRGL 399  
Db 299 WNIYIEKGFINTSDNSTRVRAGESLSLRVMVMAKPKHPTFSWSYS----- 346  
Qy 400 DGYSISKFCDH-----KKNPGEYIFVAENDDAQFTKMTNLINIRK 438  
Db 347 -GVKLTNTDHWITRTHGNSYTSKLVRLKVSSEGIYTFSCLRDATIRTFEVHVIS 405  
Qy 439 KPQVLANASA--SQASCSGDYPLPSWTW-----KKSDKSPNCTEPIPEGVWNKKAN 489  
Db 406 KPQIVSYEGPIDGQVRCVAEGYPTQIKWYCDLPHSRCSNLL-NATQE-EEDVVTVTMT 463  
Qy 490 RKVFGQWSSSTLNMSEAGKGLLVKCCAYNSMGTSCTIFLSPGFPF---FLQDNISFY 546  
Db 464 NPPFGKGAVERSLNTKNNYATLE--CVASANGEIYVTLFSISENTVPHELTPLLIGFV 521  
Qy 547 ATIGLCPLFIIVLIVLILCHYKQFRYESOLOMIOVTPLDNEIFYVDRDYEYDLKWEF 606  
Db 522 AAVI-----LVILIVLITRYKQKFKYQKQVIE--GIHGNNYVIDTQPLPYDQWEEF 575  
Qy 607 PRENLEFGKVLGSGAFGRVYNNATAYGISKTGVSIOVAVKMLKEKADSCKEALMSELKM 666  
Db 576 PRDLRFKGTGLGSGAFGKVVATAYGMSKADTVMTVAVKMLKPSAHATEKEALMSELKVL 635  
Qy 667 THLGHNDTNVNLGACTLSGPVYLFEYCCYGDLLNYLRSKREKFRHTWTEIFKEHNFSS 726  
Db 636 SYLGNHINTVNLGACTVGPTLVITEYCCFGDLLNLRRLRRVIFYFT---TLGDAYVR 692  
Qy 727 YPTFOAHSSNMPGSGREVOLHP--LDOLSGFNGNSTHSDEIEYENOKRLAEEREEDLNV 785  
Db 693 NVNMQSEPNDSRG--YMTKKFSVGLISENRRSLNKGD--SYSDSDAVSEILQEDGLT 748  
Qy 786 LTFEDLLCFAYQVAKMEFLFKSVYHRDLAARNVLVTHGKVKVICKDFGLARDILSDSSY 845  
Db 749 LQTEDELLSFYQVAKGMDFLASKNCIHRDLAARNILLTQGRVAKICDFGLARDITTDNSY 808  
Qy 846 VVRGNARLPVKNMAPSLPFGIYITIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYKLI 905  
Db 809 VVKGNARLPVKNMSPESIFECVYTFESDWSYGILLWEIFSLGSSYPGPGMPVDSKFKMI 868  
Qy 906 QSFKMEQPPYATYEGIVFMQSCWADSRKRPSPNLTSLFLGQLAFA--EACIRTSIHL 964  
Db 869 KEGYRMSEPFSPSEPMYDMHSCWDADPVKRPSPFKIVEKIQOISDSTKHYILNFSRL 928  
Qy 965 PKAAAPQQRGLRAQSPQRQVQKTHRRS 992

Db 929 P--AAP-----GPRESSH---VHRLNS 947  
  
RESULT 2  
Q63702 PRELIMINARY; PRT; 974 AA.  
ID Q63702  
AC Q63702  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE C-kit receptor tyrosine kinase isoform.  
OS Rattus rattus (Black rat)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID:10117;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BN/FM1;  
RT Tsujimura T., Tono T., Yamazaki M., Nomura S., Kitamura Y.;  
RA "Two isoforms of rat c-kit receptor tyrosine kinase."  
RL Nucleic Acids Res. 0:0-0(0).  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
CC PROTEIN KINASES  
DR EMBL; X62491; CAA44354.1; -  
DR HSSP; P11362; IFCK.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR003598; Ig\_C2.  
DR InterPro; IPR003600; Ig\_Like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR001824; RTKinaseIII.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00047; ig; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 2.  
DR SMART; SM00408; IGC2; 1.  
DR SMART; SM00410; IG\_Like; 3.  
DR SMART; SM00219; TyKIC; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
KW Phosphorylation; Receptor; Transferase; Transmembrane;  
KW Tyrosine-protein kinase.  
SQ SEQUENCE 974 AA; 108955 MW; B9BFAD9BA0DA190D CRC64;  
  
Query Match 23.5%; Score 1234.5; DB 11; Length 974;  
Best Local Similarity 32.3%; Pred. No. 1.3e-99;  
Matches 323; Conservative 176; Mismatches 340; Indels 161; Gaps 37;  
  
Qy 36 CVLISHENNGSSAGKPPSYRMVRGSPEDLQCTPRQSEGTVEAATVEAEGSITLQVQ 95  
Db 12 CVLVLVLRGQTGTSQSA-----SPGEPSPPSIOPAQSELIVEAGD-TIRLT 57  
Qy 96 LATPGDLSCLVFKHSLGQCPHFDLQNRGIVSMALNV-TETQAGEYLLHIQSERANYT 154  
Db 58 CTDPAFVK--WTEF-----ILDVRIENKQSEWIR--EKAETHT 92  
Qy 155 VLFT-----VNRD--TQLYVLRPRYFRKMKNDALLCISEGVPEPTV-EMVLC 200  
Db 93 GKYTCVSGSLRSIYVFRDPAVLFLVGLPLFKEDNDALVRC---PLTDPOQVNSYSLI 149  
Qy 201 SSHRESO-----KEGPAVVRKEEKVLHLEFGTDIRCCARNALGRECTKLTIDLNQ 252  
Db 150 EDCGKSLPTDLKVPNPKAGITIKNVRAYHRLC---IRCAAQREGKWMRSDDKFTLVKRA 206  
Qy 253 A-----POSTLPQL--FLKVGPELWIRCKAIHVNHGFLTW-----ELEDKALEE 295  
Db 207 AIKAIPLVVPETSHLLKEGDTFTVICTIKDVSTSVDSMWIKLNPQPSKQAKVKNRNSWHQ 266  
Qy 296 GSYFEMSTYTNRTMIRILLAFVSSVGRNDGYTCSSSKHPSQSALVT---ILEKGIN 352

```

Db 267 GDF-----NYERQETLT-----ISSARVNDSGVFMCMYANNNTFGSANVTTLTKVVEKGFN 316
QY 353 --ATSSQEEYIDPYEKFCFSVRFKAYPR-IRCTWIFSOASFCEQORGL-----DGYSSIS 405
Db 317 IFPVKNTTVFVTDG-ENVDLVVEFEAYPKPEHOOWIYMNRT--PTNRGEDIYKSDNOSNI 373
QY 406 KFCDH-----KNKPGEIFYFAENDDAQFTKMTLNIRKKPOVLA--NASASOASCSDD 456
Db 374 RYVNELRLRLKGTGGTYTFLVNSDSVASVTFDVIYVNTKPEILYDRLMNGRLQCVAA 433
QY 457 GYPLPSWTWKKCKSDKSPNCTEETPE-GVWNKANRKFVQGWSSSTLNMSSEAGKGLLVK 515
Db 434 GPEPTIDWYFTGABQRCTVPVPPVDVQIONASVSPFGKLVQSSDVSFRHNGTVEC 493
QY 516 CAYNSMGTSCTIFLNSPGFPF-----IQDN-----ISFYATIGLCLPFVVLIV 561
Db 494 KASNAVGS--SAFFN----FAFKGNSKQIQPHLTFTPLLTGFFVYTAGL-----MGIIVM 543
QY 562 LICHYKKQFRYESOLOMI-QVTGPLDNEFYVDFRDYEDLKWEPFRENLEFGKVLGSG 620
Db 544 VLAYKLOKPMYEVQWVVEEING---NNVYIDPTQLPYDHWKPEFNRNLSFGKTLGAG 600
QY 621 AFRGVNATAYGISTGYSIQVAVMKLEKADSCKEALMSKLMTHLGHHDNIYNLLG 680
Db 601 AFGKVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSKLVLSLGNHNMIVNLLG 660

Db 267 GDF-----NYERQETLT-----ISSARVNDSGVFMCMYANNNTFGSANVTTLTKVVEKGFN 316
QY 353 --ATSSQEEYIDPYEKFCFSVRFKAYPR-IRCTWIFSOASFCEQORGL-----DGYSSIS 405
Db 317 IFPVKNTTVFVTDG-ENVDLVVEFEAYPKPEHOOWIYMNRT--PTNRGEDIYKSDNOSNI 373
QY 406 KFCDH-----KNKPGEIFYFAENDDAQFTKMTLNIRKKPOVLA--NASASOASCSDD 456
Db 374 RYVNELRLRLKGTGGTYTFLVNSDSVASVTFDVIYVNTKPEILYDRLMNGRLQCVAA 433
QY 457 GYPLPSWTWKKCKSDKSPNCTEETPE-GVWNKANRKFVQGWSSSTLNMSSEAGKGLLVK 515
Db 434 GPEPTIDWYFTGABQRCTVPVPPVDVQIONASVSPFGKLVQSSDVSFRHNGTVEC 493
QY 516 CAYNSMGTSCTIFLNSPGFPF-----IQDN-----ISFYATIGLCLPFVVLIV 561
Db 494 KASNAVGS--SAFFN----FAFKGNSKQIQPHLTFTPLLTGFFVYTAGL-----MGIIVM 543
QY 562 LICHYKKQFRYESOLOMI-QVTGPLDNEFYVDFRDYEDLKWEPFRENLEFGKVLGSG 620
Db 544 VLAYKLOKPMYEVQWVVEEING---NNVYIDPTQLPYDHWKPEFNRNLSFGKTLGAG 600
QY 621 AFRGVNATAYGISTGYSIQVAVMKLEKADSCKEALMSKLMTHLGHHDNIYNLLG 680
Db 601 AFGKVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSKLVLSLGNHNMIVNLLG 660

```

```

DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001824; RTKinaseIII.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; ig; 1.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 2.
DR SMART: SM00408; IGC2; 1.
DR SMART: SM00410; IG_like; 3.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane;
KW Tyrosine-protein kinase.
KW
SQ SEQUENCE 978 AA; 109341 MW; 0958C33F19889051 CRC64;

```

```

Query Match 23.4%; Score 1232.5; DB 11; Length 978;
Best Local Similarity 32.2%; Pred. No. 2e-99;
Matches 323; Conservative 176; Mismatches 340; Indels 165; Gaps 37;

QY 36 CVLISHENNGSSAGKPSYRMVRGSPEDLOCTPRRQSEGTYYEATVEAESGSIITLQV 95
Db 12 CVLLVLLRGOTG*SQPSA-----SPGEPSPPSIQPAQSELIVEAGD-TIRLT 57
QY 96 LATPGDLSCLWPKHSLGCPHFEDLONRGIVSMAILNV-TETQAGEYLLHIQSERANYT 154
Db 58 CTDPAFVK--WTFE-----ILDVRIENKQSEWIR--EKAEATH 92
QY 155 VLFT-----VNVRD-TQYVLRPRYFKRMNQDALLCISGEVPEPTV-ENVLC 200
Db 93 GKTYVSGSLRSIIYFVDPDAVFLVGLPLFGKENDALVRC---PLTDPOVSNVSLI 149
QY 201 SSHRESC-----KEGPAVVRKEEVHLHFGTDIRCCARNALGRECTKLTIDLNQ 252
Db 150 ECDGKSLPTDLKFPNPKAGITIKNVKRAYHRLC---IRCAAQREGKWMRSDFKTLAVRA 206
QY 253 A---PQSTLPOL--FLKVGPELWIRCKAIHVNHGGLTW-----ELEDKALEE 295
Db 207 AIKAIPIVSVPTSHLLKEGDTFTVICTIKDVSIVSDSMWIKLNPOQSKAQKRNWHQ 266
QY 296 GSYFEMSTYSTNRTMIRILLAFVSSVGRNDTGYTCTSSSKHPSQSALVT---ILEKGFN 352
Db 267 GDF-----NYERQETLT-----ISSARVNDSGVFMCMYANNNTFGSANVTTLTKVVEKGFN 316
QY 353 --ATSSQEEYIDPYEKFCFSVRFKAYPR-IRCTWIFSOASFCEQORGL-----DGYSSIS 405
Db 317 IFPVKNTTVFVTDG-ENVDLVVEFEAYPKPEHOOWIYMNRT--PTNRGEDIYKSDNOSNI 373
QY 406 KFCDH-----KNKPGEIFYFAENDDAQFTKMTLNIRKKPOVLA--NASASOASCSDD 456
Db 374 RYVNELRLRLKGTGGTYTFLVNSDSVASVTFDVIYVNTKPEILYDRLMNGRLQCVAA 433
QY 457 GYPLPSWTWKKCKSDKSPNCTEETPE-GVWNKANRKFVQGWSSSTLNMSSEAGKGLLVK 515
Db 434 GPEPTIDWYFTGABQRCTVPVPPVDVQIONASVSPFGKLVQSSDVSFRHNGTVEC 493
QY 516 CAYNSMGTSCTIFLNSPGFPF-----IQDN-----ISFYATIGLCLPFVVLIV 561
Db 494 KASNAVGS--SAFFN----FAFKGNSKQIQPHLTFTPLLTGFFVYTAGL-----MGIIVM 543
QY 562 LICHYKKQFRYESOLOMI-QVTGPLDNEFYVDFRDYEDLKWEPFRENLEFGKVLGSG 620
Db 544 VLAYKLOKPMYEVQWVVEEING---NNVYIDPTQLPYDHWKPEFNRNLSFGKTLGAG 600
QY 621 AFRGVNATAYGISTGYSIQVAVMKLEKADSCKEALMSKLMTHLGHHDNIYNLLG 680
Db 601 AFGKVEATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSKLVLSLGNHNMIVNLLG 660

```

```

RESULT 3
Q63116 ID Q63116 PRELIMINARY; PRT; 978 AA.
AC Q63116;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE C-kit receptor tyrosine kinase.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RX MEDLINE=92003944; PubMed=1912577;
RA Tsujimura T., Hirota S., Nomura S., Niwa Y., Yamazaki M., Tono T.,
RA Morii E., Kim H., Kondo K., Nishimune Y., Kitamura Y.;
RT "Characterization of ws mutant allele of rats: A 12-base deletion in
RT tyrosine kinase domain of c-kit gene.";
RL Blood 78:1942-1946(1991).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL: D12524; BAA02094.1; -.
DR HSSP; P11362; IFGK.

```

QY 681 ACTLSPVYLIFECYCGDLLNLYRSKREKPHRTWTEIFKEHNFSSYPTFOAHNSMMPG 740  
 DB 661 ACTVGPVILITEYCCYCGDLLNLYRSKREKPHRTWTEIFKEHNFSSYPTFOAHNSMMPG 718  
 QY 741 SREVOLHPDLQSGFN-----GNSIHSEDEIEYENQKRLAEEDLNVLTFEDL 791  
 DB 719 SNEY-----MDMKPGVYVVPKTKDKRSARSDSYIERDVTVAIMEDDELADL---EDL 770  
 QY 792 LCFAYQVAKGMFLEFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDILSDSSVYVRGNA 851  
 DB 771 LSFQVQVAKGMFLEFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDILSDSSVYVRGNA 830  
 QY 852 RLPVKMAPESIFEGYITKSDVMSYGLLWELFSLGVPNDYGPVDFANFYKLIQSGFRM 911  
 DB 831 RLPVKMAPESIFNCVYTESDWSYGLLWELFSLGVPNDYGPVDFANFYKLIQSGFRM 890  
 QY 912 EQPFYATEGIYFVQSCWAFDSKRPSPNLTSLFLCQLAEAE 955  
 DB 891 LSPHEAPAAAYEMKTCWDADPLKRPFKOVVOLIEKQISDSSK 934

RESULT 4  
 Q91909 PRELIMINARY: PRT; 954 AA.  
 AC Q91909;  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE C-kit-related kinase 1 (Akr1) precursor.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_Taxid=8355;  
 RN [1]  
 RZ SEQUENCE FROM N.A.  
 RX MEDLINE=9534496; PubMed=7619732;  
 RA Baker C.V., Sharpe C.R., Torpey N.P., Heasman J., Wylie C.C.;  
 RT "A Xenopus c-kit-related receptor tyrosine kinase expressed in  
 RT migrating stem cells of the lateral line system";  
 RL Mech. Dev. 50:217-228(1995).  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 DR EMBL: Z48770; CAA88688.1; -;  
 DR HSP; P11362; IFGK.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR003600; Ig-like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR01824; RTKinaseII.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00047; Ig; 2.  
 DR Pfam: PF00069; pkinase; 1.  
 DR ProDom: PD000001; Euk\_pkinase; 2.  
 DR SMART: SM00410; IG\_like; 2.  
 DR SMART: SM00219; Tyrc; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
 KW Phosphorylation; Receptor; Signal; Transferase; Transmembrane;  
 KW Tyrosine-protein kinase.  
 FT SIGNAL 19  
 KW POTENTIAL.  
 SQ SEQUENCE 954 AA; 106859 MW; 9748845CBE0B537D CRC64;

Query Match 32.2%; Score 1220; DB 13; Length 954;  
 Best Local Similarity 32.0%; Pred. No. 2.4e-98;  
 Matches 315; Conservative 160; Mismatches 321; Indels 188; Gaps 35;  
 QY 77 YEAAATVEAESGIFLQVLAATPGDLSL-----LWFKHSSLGCPQHFQDQNRGIVSMA 130  
 DB 15 YTGDAVPKINDEDRVTNVGDKVSLCERDAHLVTLATQKSLGMLKPKR-DLKSRLP----- 68

QY 131 ILNVITET-----QAGEYLLHIOSEANVTLVFTVNRDQTLVLRPPYPRKME 178  
 DB 69 -LNNSETDOFFVIKADLRHIGRYICTNTQENTSV--SLFKDPAARFLDIPFDIVTE 125  
 QY 179 NODAL-LCISEGVEPTVEMWLCSSHRESCKEGPAVVRKEEKVLHELFGTDIRC----- 232  
 DB 126 GADTVGMCFP---TDPMDIAI-----EKC--DGSPLPE-----NFTTDTIEAGITIK 169  
 QY 233 -----CARNALGR-ECTKLTIDLNQAPQSTLQQLFL-----KVCEPLWIRC 273  
 DB 170 TVOLAFDSCYVCGSGNKTGVRKSSSTSIHVKVPKK-VPTVFLSKSRQLVKTEPEVTC 228  
 QY 274 KAIHVNHGGLTW-ELEDKALEBSYFENSTYSTNRTMIRILLAFVSSVGRNDTGYTCS 332  
 DB 229 AVLDFESTVKAOWLDVKEGVTQANFRSSNVFSYNLTLS-----DGVPSRSRFTCQ 282  
 QY 333 SKKHPQ---SALVTILEKGFINATSOE-EYEIDPYEKFCFSVRFKAYPR-IRCTWIFS 387  
 DB 283 AENAIGQVNAFTLDVIDVGVNLTLENTTISVNAAGDNLVAVYIDAYPHDPDGVWTFY 342  
 QY 388 QASFPCEQRLGEGYSISKFDHKN-----KPEYIFYAENDDAQFTKMF 432  
 DB 343 NETL---LNTSDHYVATK--DEGNRRYSELHLIRLKGTEKGVYTFYTTNSDDASVSF 396  
 QY 433 TLNIRKKPOVLANASASQAS--CSSDGYPLPSWTWKKCDKSPNCTEEIPEGVWKK--A 488  
 DB 397 NIOVKTRPEILIAERTSEGTLQCVATGFPVPAIOWEFCPGSEQRCTDYPPLSPVNEFIQ 456  
 QY 489 NRKVFQGWSSSTLNMSEAGKLLVKKCAVSMGTS-----CETFLNSPGP 535  
 DB 457 ENSLGRVIVESTIDVNDLKKNGTVQCVASNEVASVFSFAIKELRTHLTFT----- 511  
 QY 536 FPFIDQNIISFYATIGLCLPFIVLVILVILCHYKQFYESOLQMI-QVTGPDNEYFVVD 594  
 DB 512 -PLL---IGFIAAAGL---MCIATAVAVLMVKYLOKPKYEQWKVVEEING---NNYVID 560  
 QY 595 FRDYEDLKWEFFRENLEFGKVLGSGAFGRVMNATAGISKTGVSIOVAVKMLKEKADSC 654  
 DB 561 PTQLPYDNKWEFFPRDLRCFGKILGAGAGFGVVEATAYGLLKEDSRLTAVAVKMLKPSAUST 620  
 QY 655 EKEALMSLKMTHLGHHDNIVNLLGACTGSPVYLIFECYCGDLLNLYRSKREKPHRT 714  
 DB 621 EREALMSLKVSLYGLHHKNVNLGACTGSPVYLIFECYCGDLLNLYRSKREKPHRT 677  
 QY 715 WTEIFKEHNFESSYPTFOAHNSMMPGSEVOLHPDLQSGFNHSHSEDEIEYENOK- 773  
 DB 678 -----ICPKFEDNS-----EAALYKNL-----LNRDMGCEGMSEYIDHKP 713  
 QY 774 -----RLAEDEEDLNVLTFEDLFCFAYQVAKGMFLEFKSC 810  
 DB 714 AVSYVVPKTKDKRSRSGSGFDQSVSIPEEDDL-ALDTEDLINFSYQVAGNMLASKNC 772  
 QY 811 VHRDLAARNVLVTHGKVKIKCDFGLARDILSDSSVYVRGNAARLPVKMAPESLFGGIYTI 870  
 DB 773 IHRDLAARNVLVTHGKVKIKCDFGLARDILSDSSVYVRGNAARLPVKMAPESLFGGIYTI 832  
 QY 871 KSDVMSYGLLWELFSLGVPNDYGPVDFANFYKLIQSGFKMEQPFYATGIVFVMSQSWA 930  
 DB 833 ESDVMSYGLLWELFSLGVPNDYGPVDFANFYKLIQSGFKMEQPFYATGIVFVMSQSWA 892  
 QY 931 FDSRRKRPSPNLTSLFLCQLAEAE 954  
 DB 893 SDPLKRPSPNLTSLFLCQLAEAE 916  
 RESULT 5  
 Q98SU3 PRELIMINARY: PRT; 974 AA.  
 ID Q98SU3;  
 AC Q98SU3;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)

DE Receptor tyrosine kinase Fms.  
GN FMS.  
OS Danio dangila.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=127599;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21359118; PubMed=11466528;  
RA Parichy D.M., Johnson S.L.;  
RT "zebrafish hybrids suggest genetic mechanisms for pigment pattern  
RT diversification in Danio.";  
RL Dev. Genes Evol. 211:319-328(2001).  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
CC PROTEIN KINASES.  
DR EMBL; AF324479; AAK15301.1; -.  
DR HSP; P11362; 1FGK.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR003600; Ig\_like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR001824; RTKinaseIII.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00047; Ig\_3.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 2.  
DR SMART; SM00409; IG; 4.  
DR SMART; SM00408; IGC2; 1.  
DR SMART; SM00410; IG\_like; 1.  
DR SMART; SM00220; S-Trc; 1.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
KW Phosphorylation; Receptor; Transferrase; Transmembrane.  
FT VARIANT 720 720 A -> T.  
SQ SEQUENCE 974 AA; 109902 MW; DE7FD64FD0D78B37 CRC64;

Query Match 23.1%; Score 1215.5; DB 13; Length 974;  
Best Local Similarity 34.6%; Pred. No. 6.3e-98;  
Matches 339; Conservative 147; Mismatches 342; Indels 153; Gaps 37;

QY 54 YMWVRGSPEDLQCTPR-RQSEGTVEAATVEAEGSITLOVOLATPGDLSCLWFKHSS 112  
Db 13 EQVOGWSE-----PRILSSGAL--ADTDVLDGSGS---PLQLCEGGPVTFL----- 57

QY 113 LGCQPHFDLQNGRI-----VSMALNVETQAGEY-LLHIQSERANVTVLFTYN 160  
Db 58 ----PRLAKHKRYISKEVGKSTRFRKATVDFT---GTKCIYINGNDSNLSSSVHVF 109

QY 161 VPDQO-LIV-----LRRYFRKMNQDALLCISEGVPTVEWLCSSHRESCKEENAV 214  
Db 110 VNSRVLFVSPSTSLR--YVRK-EGEDLLLPCLLDPPAT-DFIFRMONGSAAPYGMNAT 165

QY 215 VRKEKVL-----HELFGTDIRCCARNALGRECTKFTLDLNAQPOSTLPQLFLK----- 264  
Db 166 FDPKGVLRNVHPGNDYICSAIGGAERKSKIFSINVIRLRFP-PYVYLKRENYVK 224

QY 265 -VGEPLWIRKAIHNVGHGFTWELEDKALEGYSFYFENSTYNTMTIRILLAFYSSVGR 323  
Db 225 LVGERLQISCTNNFNFSYNTVTHSSRLPAE--ERSTMEDGLATESILT-IPSVQL 281

QY 324 NDTGYTCSSKHPQSQALVT---ILEKGF-----NATSSQEEYIDPYERFCFSV 372  
Db 282 SDTGNTCTGNEAGANSSTOLLVDEPYIRLSPKLSKLTNRDLSIEVSGDDVDLGV 341

QY 373 RFKAY-PIRITW----IFSQASFCEQR--GLEDGYSISKFDHKN--KPGEYIFYAEND 424

Db 342 WICAYPLPLSHKWTPTSHNASLP-ENRFYNHNDRIEALLFLKRLNFEEICQYTLNVKNR 400  
QY 425 DAOFTKMTFTNIRKKQVLAN-ASASQASGSSDGYPLPSTWTKKCSOKSPNCTEE----- 478  
Db 401 VKSASITFDIKMYTKPVARVKWENVTTLSRSYGYPAISILWYQCTGIRTTCPENTDLO 460  
QY 479 -----IPEGVNNKKANRVFGQWSSSTLNMSSEAGKGLLVCCAYNSMTSCETIFLN 531  
Db 461 PIQTHGVPGE-----ESFGVGVESVLTVG-PNRRMTVVCVARNLVGGSDTFSMD 511  
QY 532 SPGPFPIQDNISFYATIGLCLFIV--VLIVLICHYKKOFRYESQLOMQVTGPLDNE 589  
Db 512 -----VSQDITSACMGSTVAMVGLLLIFMIYIKQKPRYEIRWKIIEATN--GNN 562  
QY 590 YFYVDRDYEDLKWEPRENLEFGVLGSGAFGRVNNATAYAGISKTGVSIOQAVKMLKE 649  
Db 563 YTFIDPTQLPYNEKWEPPRDKLKLKTLGAGAGKVVYATAYGLGKEDNITRVAVKMLKA 622  
QY 650 KADSCKEALMSLKMTHLGHNDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYRSKRE 709  
Db 623 SAHPDEREALMSLKLHLGQHKNIIVNLLGACTHGGPVLVITEYCCCHGDLNLSKAE 682  
QY 710 KEHRTWTEI--FKE-----HNFSSYPTFOAHSSNMPGS-----REVQLHPPLDLS 754  
Db 683 NLFUNVMTIPDFPEPMTDYKNISTERMF-VRSDSGISSACSDHYLDMRPVTSRPTLSSS 741  
QY 755 GFNGNSIHSEDEYEYENQKRLAEEDNLVTFEDLLCFAYOVAKGMFELEKSCVHRD 814  
Db 742 -----ECQEDSWPLDMDDLLRFSQVAAQGLDFLAANKCHRD 778  
QY 815 LAARNVLVTHGKVKYKICDFGLARDILSDSYVVRGNARLPVKWMAPESEFEGYITIKSDV 874  
Db 779 VAARNVLLTNSRVAKICDFGLARDIMDSNVYVKGARLPVKWMAPESEFECYVTVQSDV 838  
QY 875 WSGVILLWEIFSLGVNYPGIPVDANFYKLIQSGFKMEQPFYATEGYFVWQSWADSR 934  
Db 839 WSYGIMLWEIFSLGKSPYPNILDYSKFYKMKCYQMSRDPFAPPENYTIMKMCNLDAA 898  
QY 935 KRPSFNLTSFLGCQLAAEAE 955  
Db 899 ERPTFSKISQIMORMLGETPE 919

RESULT 6  
Q98SU1 ID Q98SU1 PRELIMINARY; PRT; 977 AA.  
AC Q98SU1;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created).  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Receptor tyrosine kinase Fms.  
OS Danio nigrofasciatus.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=144739;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21359118; PubMed=11466528;  
RA Parichy D.M., Johnson S.L.;  
RT "zebrafish hybrids suggest genetic mechanisms for pigment pattern  
RT diversification in Danio.";  
RL Dev. Genes Evol. 211:319-328(2001).  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
CC PROTEIN KINASES.  
DR EMBL; AF324481; AAK15303.1; -.  
DR HSP; P11362; 1FGK.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR003600; Ig\_like.



```
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001824; RTKinaseIII.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00047; Ig_4.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 2.
DR SMART: SM00409; IG; 4.
DR SMART: SM00408; IGC2; 1.
DR SMART: SM00410; IG-like; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
DR ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transmembrane.
FT VARIANT 60 V -> L.
FT VARIANT 103 103 A -> S.
FT VARIANT 742 742 P -> S.
FT VARIANT 899 899 E -> D.
SQ SEQUENCE 977 AA; 109881 MW; CF54E129FAB10E3E CRC64;

Query Match 23.1%; Score 1215.5; DB 13; Length 977;
Best Local Similarity 33.5%; Pred. No. 63e-98;
Matches 334; Conservative 160; Mismatches 356; Indels 147; Gaps 35;

QY 54 YRMVRSPELQCTPR-ROSEGTVYEAATVEAESITLQVOLATPGDLSCLW---VFK 109
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 13 FGQVQGWSE-----PRILNSGAL--AGTDVILDSGS--PLQLVCGEGDGPVTFPRVAK 62

QY 110 HSLGCPHFQDLNQRGIVSMALINVTETONGEY-LLHQISERANYTVLFTVNVRTQ-LY 167
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 63 H-----KRYISKEVGIRSFVEKATVDFGTYKCVINGSNLSASSVHFVRKSVLF 117

QY 168 V-----LRPRFKMENQDALLCISEGPEPTVEWLVLCSSHRESCKEGBGPAVVRKEKVL 222
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 118 VSPSSSLR--YVRK-EGEDLLLCLLTDPDAT--DFTFRMDNGSAAPYGMNATFDPKGV 173

QY 223 ----HELFGTDIRCCANLGRCTKFTIDLNOAPOSTLPQLFLK-----VGEPLWI 271
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 174 IRNVHGFNADYICSAIGCAKGVSKIFSINVIOIRLRF--PYVYLKNEYVVKLVGERLQI 232

QY 272 RKAHIVNHGFGLTWELEKALBEGSVFEMSTYSTNTMIRILLAFVSSVGRNDGTGYTC 331
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 233 SCTTNPNFNYVTWTHSSKRLPKPE--EKSTMEGDRLAIESILT-ISSVQLSDTGNTIC 289

QY 332 SSSKHPQSALVT---TLEKGFINAT---SSQEEY-----EIDPYERKCFSVRFKAYPRI 380
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 290 TGONEAGANSSTQLLVVDEPYIRLSKLSKLTGRLSIEVSEGEDVDLGLVLEAYPPL 349

QY 381 RC-TW----IFSQASFPCEQR--GLEQGYISKFCODHKN--KPGEYIFYAENDDAQTKMF 432
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 350 TSHWTETPTSHNASLP--ENRFYHNDRYEALLFLKLNFEIEIGQYTLNVKNSMKSASITF 408

QY 433 TLNIRKPKQVLAN--ASASQSCSSDGPPLPSWTHKKCKSDKSPNCTEEI----PEGVNKK 487
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 409 DIKMYTKPVARVKNWNTTILSCRSYGYPAISILWYQCTGINTCPENTDLOIQTQVTE 468

QY 488 ANRVFGQWSSSTPLNNSKAGKGLLVKCCAVNSMGTSETIFLSPGPFPIQDNISFYA 547
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 469 FQKESFAGVGVESVLTVG--PNRRMTVVVCVAFNLVQGSDFTEFSD-----VSQIIFSA 520

QY 548 TIGLCLPFIV--VLIVLICHKKYKOFRESQLOMIQVTPGLDNEYFYVDFRDYEYDLKWE 605
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 521 MCGSTVAMVVLGILLIFMIYKQKPRYERIKWKIEATN--GNNYTFIDPTQLPYNEKWE 578

QY 606 PPENLEFGKVLGSGAFGRVNMATAYGLSKTGVSVIQAVKMLKEKADSCKEALMSLKM 665
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 579 FPRDKLGLKTLGAGAFGKVVYEATAYGLGKEDNITRVAVKMLKASHPDEREALMSLKI 638

QY 566 MTHLGHHDNINVLGACTLSGPVLIFFEYCCYGDLLNLRKREKF----- 711
```

```
Db 639 LSHLGQHKINVLNLLGACTHGGPVLVITEYCCCHGDLNLFLSKAEENLFNFTIPNPEPV 698
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 712 -----HRTW-----TEIFKEHNFSSYPTFOAHSSNMSPGSRREVOLHPPDLQLSGF 756
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 699 TDYKNVSTERMFVRSDSGISSTCSHDYLDMPRTVSRPTNSALDPSSDCQ----- 747
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 757 NGSNIHSEDEIEYENOKRLAEEDLNVLTFEDLLCFAYQVAKGMEFLFEKSCVHRDLA 816
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 748 -----EDSWPLDMDLDRFSSQVAGGLDFLAAKNCIHRDVA 783
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 817 ARNVLTGHKVKVVICDFGLARDILSDSSVYVVRGNARLPVKWMAPESLFEGIYTIKSDVWS 876
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 784 ARNVLTNSRVAKICDFGLARDIMNSNYYVVGKGNARLPVKWMAPESEFECVYTVQSDVWS 843
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 877 YGILLWEIFSLGVNYPGPVPDANFYKLIQSGFKMEOPFYATEGIYFVMOSCAFDGRKR 936
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 844 YGIMLWEISLGKSPYNILVDSFKYKMIKCGYQMSRPDFAPEPMYTIMKMCNLEAAER 903
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 937 PSFPNLTFLGCOLAE-AEEACITSIHLPKQAAPQ 972
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 904 PTFKSIQMIQRLGETSEQDQTEYKNIPSEAVEQ 940
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
QYI8N6 PRELIMINARY; PRT; 977 AA.
AC Q9I8N6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Fms.
GN CSF1R OR FMS.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB.
RX MEDLINE=20323166; PubMed=10862741;
RA Parichy D.M., Ransom D.G., Paw B., Zon L.I., Johnson S.L.;
RT "An orthologue of the kit-related gene fms is required for development
of neural crest-derived xanthophores and a subpopulation of adult
melanocytes in the zebrafish, Danio rerio."
RL Development 127:3031-3044(2000).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
PROTEIN KINASES.
CC EMBL: AF240639; AAF76872.1; -.
DR HSXP; P11362; IFGK.
DR ZFIN; ZDB-GENE-001205-1; csf1r.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001824; RTKinaseIII.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig_4.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR ATP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
KW Receptor; Transmembrane; Tyrosine-protein kinase.
FT VARIANT 238 238 N -> K.
FT VARIANT 615 615 V -> M.
SQ SEQUENCE 977 AA; 110187 MW; C91A2F339E746A58 CRC64;
```

Query Match	23.1%	Score 1214;	DB 13;	Length 977;
Best Local Similarity	33.4%;	Prod. No. 8.6e-98;		
Matches	329;	Conservative 157;	Mismatches 354;	Indels 146; Gaps
Qy	68	PR-RQSEGTVEAATVEAESGSITLQVLATPGDLSCLW---VFKHSSLCGPHFDLQN 123		
Db	22	PRILSSGAL--AGTDVILESGS---PLQLVCEGDGPVTEFLPKAH-----KRYISKEV 71		
Qy	124	RGIYSMAILNVTQTAGEY-LLHIQSRANVTVLFTVNVRDTO-LVY-----LRRPYPERK 176		
Db	72	GKIRSFREKTTVDFTGYKCVYMGNGDNSLSSVHVFDSDRVLFVSPSTSLR--YVRK 129		
Qy	177	MENODALL-CISEGVPBPTVEWVLCSSHRECKBEEGPAVVRKEEK-----VLHELFCTDI 230		
Db	130	-EGEDLLLPCL--LTDPEATDFTFRMDNGSAAPYGMNITYDPKGVGLIRNVHPGENADY 185		
Qy	231	CCARNALGRECTKLFITDLNQAPOSTLPOLFLK-----VGEPLWIRCKATHVNHGEG 283		
Db	186	ICCARIGAEKVKISFINIIQRFPF-PVYILKRNEYKVLGVERLQISCTTNPNPFYIN 244		
Qy	284	LTWELEDKALEEGSYFEMSTYTNRTMIRILLAFVSSVGRNDTGYTCSSSKHPSOSALV 343		
Db	245	VTWTHSSRMLPKAE--EKSTWEGDRLAIESILT-IPSVOLSHTGNTCTCGONPAGANSST 301		
Qy	344	T---ILBKGFINAT---SSOEY-----EIDPYKEFCFSYRFKAYPRIRC---TWIFSQ 388		
Db	302	TQLLVAPEPVIIRLSPKLSSKLTHRGLSIEVSEGDVDLGLVIEAPPLTSKHWEPTPTSHN 361		
Qy	389	ASPPCEQR--GLENGYISIKFCDBKN--KPEYIFFAENDDAOFTKMFILNIRKKPOVILA 444		
Db	362	ASLP-ENRFTNHNDRYEALLLLKRLNPEEIGQYTLNVKNSMKASITFDIKMYTKPVARV 420		
Qy	445	N-ASASQASDCSDGYPLPSPWTKKSDKSPNCTEEI---PEGWNKANKRVFGQVSS 499		
Db	421	KWENTVTLSCRSYGYPAPSLNLTQCTGIRTCENTDLOPIQTQVFQKESFGAAGVE 480		
Qy	500	STLMSAAGKGLLVKCAYNMGTSCETIFLNSPGPPFIQDNISFYATIGLCLPPIV---557		
Db	481	SVLTVG--PNRRMTVVCVAFNLVGOGSDTFSME-----VSDQIFTSAMCGSTVAMVLG 532		
Qy	558	VLIVLICHKKYKQFRYESQLOMQVTPGLDNEYFYVDFRDYEDLKWEPRENLEPGKVL 617		
Db	533	LLLIIFYIKYKQPRYEIRKWIIEATN--GNNTYTFDPTQLPYNEKWEPPROKLLGKTL 590		
Qy	618	GSAGFRVMNATYGISKTGVSQVAVKMLKEKADSCKEALMSELMKMTLGHHDNIYN 677		
Db	591	GAGAFKGVETATYGLCKEDNITRVAVKMLKASAHDPDEALMSEULIUSHLQGHKNIN 650		
Qy	678	LLGACTLSGPVYLIFYCYGGLNLTDLRSKRF-----HRTW 715		
Db	651	LLGACTHGGPVLVITEYCHGDLNLFRLSKAENFLNFMVTPNFPPEPMDYKNVSTERMF 710		
Qy	716	-----TELFKEHNTSSYPTFOAHNSNMSGPSEVOLHPPLDQLSGFNGNSIHSEDEIE 768		
Db	711	VRSDSGISSTCSHYLDMRPVTSRPTNSALDSSEQC-----747		
Qy	769	YENOKRLAEEBEDNLVLFEDLLFCAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKV 828		
Db	748	-----EDSPLMDMDLLRESSQVAGQLDFLAAKNCIHRDVAARNVLNLSRVA 795		
Qy	829	KICDFGLARDILSDSSVYRGNARLPVKWAPESLPEGITYTIKSDVWSYGILLWEIFSG 888		
Db	796	KICDFGLARDIMDSNVYVKGARLPVKWAPESIEFCVTVVOSDVWSYGIMLWEIFSG 855		
Qy	889	VNPYPGIPVDANFYKLIQSGKMEQPFYATEGIFYFWQSCWAFDSKRKPSFNLTSLGCG 948		
Db	856	KSPYPNILLVDSKFKMKGQYQMSRDPDAPPENYTIMKMCNWLDAERPTFSKISQIMOR 915		
Qy	949	QLAE-AEEACIRTSIHLPKQAAPQOR 973		
Db	916	MLGETSEQDDTOEYKNIPTAEAAEQ 941		

Q8WN23	PRELIMINARY;	PRT;	979 AA.
AC	Q8WN23;		
DC	01-MAR-2002 (TrEMBLrel. 20, Created)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	C-Kit.		
OS	Canis familiaris (Dog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OX	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
NCBI_TaxID=9615;			
FN	[1]		
RP	SEQUENCE FROM N.A.		
RA	zenke D., Yuzbasiyan-Gurkan V.;		
RL	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF481448; AAL40833.1;		
DR	InterPro; IPR000719; Euk_pkinase.		
DR	InterPro; IPR003599; Ig_like.		
DR	InterPro; IPR003600; Ig_like.		
DR	InterPro; IPR003006; Ig_MHC.		
DR	InterPro; IPR001824; RTKkinaseIII.		
DR	InterPro; IPR002290; Ser_thr_pkinase.		
DR	InterPro; IPR001245; Tyr_pkinase.		
DR	Pfam; PF00047; Ig_2.		
DR	Pfam; PF00069; pkinase; 1.		
DR	ProDom; PD000001; Euk_pkinase; 2.		
DR	SMART; SM00409; IG; 4.		
DR	SMART; SM00410; IG_like; 2.		
DR	SMART; SM00220; S_TKC; 1.		
DR	SMART; SM00219; TYR_KC; 1.		
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.		
DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.		
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.		
DR	PROSITE; PS00240; RECEPTOR_TYR_KIN_III; UNKNOWN_1.		
SQ	SEQUENCE 979 AA; 109753 Mw; 46C30D5DE8E33D3 CRC64;		
Query Match	23.08;	Score 1209.5;	DB 6; Length 979;
Best Local Similarity	31.38;	Pred. No. 2.1e-97;	
Matches 332;	Conservative 165;	Mismatches 369;	Indels 195; Gaps
QY	14	LLVLSVMILETVNODPLVIKCVLISHENNGSAGKPPSSVMYRGSPEDLQCTPRROSE 73	
DB	11	LCVLLLLLLGVGGSGSPSV-----SPGEPSL-----38	
QY	74	GTVYEATVVAESGSITLQVLATPGDLSCLVFKVHSSIGCGPHOLFQNGRIVSMALN 138	
DB	39	PSIHPAKSELIVSGD-ELRLSCTDPGFVK--WTFE--TLG-QLNENTHNEWITEKA--- 89	
QY	134	VTETOAGEYLHHIOSEANVTVLFTVNRD-TOLYVLRPRYPERKMNODALLCISEGVE 198	
DB	90	-----EAGHTGNYCTNRDGLRSYIVVRPAKFLVLDPLYGEGNDTLVRC----- 138	
QY	193	PTVEWLCCSHRESCKEGPAVVRKEKVLHFLGTDIR-----C--CARNALGRE 241	
DB	139	PLTDPVETNYSLRCC--EGKPLPKDLTFVADPRAGITIRNVKREYHRLCLHCSADQKGT 196	
QY	242	-CTKLFTIDLNAPOS-----TLPQLFLKVGPEPLTRCKAIHNVHGFGTLWELEKALE 294	
DB	197	VLSKKFTLKVRAAIRAPVVSVKSTSSLLKEGEAFVCMCFIKDYSFVDSFMWIKENSQOT 256	
QY	295	EGSYFEMSTYSTNRTMLRILLAFVSSVGRNDGTGYTCCSSKHPSQSALVTILE---KGFI 351	
DB	257	NAQTOSNWHHGDFNFERQEKLLISSARVNDGSMFCYANNTFGSANVTTLLEVVDKGF 316	
QY	352	NA-TSSOEEYEIDPYEKFCPSVRFKAYPR-IRCTWIFPSQASF--PCBQRLEDGYSTSKF 407	
DB	317	NIPPMSTTFIVNDGENVDLIVEAYPKPEHQOVIYMNFTFTDKWEDYKPSDNESIRY 376	
QY	408	CDH-----KNKPGEYIFAEVNDNDQAQFTKMTLNTIRKKKQVLIANASAOA--SCSSDGY 458	
DB	377	VSELHLTRLKNGEGGTYTFVOSNDVSNSSYTFVNYVNTKPEILTHESLTNGMLQCVAVG 436	

```
QY 459 PLPSWTWKCKSDKSPNCTEEI-PEGVWNNKANRKFVGOWVSSSTLNMSEAGKGLLVKCA 517
Db 437 PEPADVWTFPCPAEQRCSPVIGPMQVQMSLSLSPGKLVQSSDIYSFAKHNGTVECR 496
QY 518 YNSMTGSCETIFLNSPGPFPIQDN-----ISPYATIGLCPLPIVVLIVLI 563
Db 497 YNNVGRS--SAFFN---FAFKGNSKEQIHPTLTFTPLLIGFVIAAGM---MCIIVMIL 546
QY 564 CHKYKQFRIESQLQMI-QVTGPLDNEYFYVDFRDYEDLKWEEFRENLEFGKVLGSGAF 622
Db 547 TYKYLQKPMYEVQWKEEING---NNYVIIDPTQLPYDHKWEFFPNRLSFGKLTGAGAF 603
QY 623 GRVMTATYGISKTVGSIOVAVKMLKEKADSCKEALMSKLMTHLGHHDNIVNLLGAC 682
Db 604 GKVEATAGLLKSDAMTVAVKMLKPSAHLTEREALMSKLVSLYLGHHMIVNLLGAC 663
QY 683 TISGPLYLIFCYCCYGLDLNLYLRSRKEFRHTWTWTEIFKEHNFSSYPTFOAHSSNPGSR 742
Db 664 TVGGPTLVITECCYGLDLNLYLRSRKEFRHTWTWTEIFKEHNFSSYPTFOAHSSNPGSR 701
QY 743 EVOLHPPDLQSLGFGNNGSTHSEDEI-----EYENOK-----RLAEE 779
Db 702 EVALYK-----NLLHSESSCSDSTNEYMDMKPGVSYVYPTKADKRRSARIGSYI 751
QY 780 EEDLN-----VLTFFEDLLCFAYQAVKMEFLEKSCVHRDLAARNVLYTHGKVVKI 830
Db 752 ERDVTAPINWEDDELADLELDSFSYQVAGMAFLASKNCIHRDLAARNVLYTHGKVVKI 811
QY 831 CDFGLARDILSSSYVVRGNARLPVWMAPELSEFEGITIKSDVWSYGILLWEIFSLGVN 890
Db 812 CDFGLARDIKNSYVVRGNARLPVWMAPELSEFEGITIKSDVWSYGILLWEIFSLGVN 871
QY 891 PYGPVDFANFYKLIOSGKMEOPFYATGTYEVMOSWAFDGRKRPSNLTSLGCOL 950
Db 872 PYGPVDFANFYKLIOSGKMEOPFYATGTYEVMOSWAFDGRKRPSNLTSLGCOL 931
QY 951 ARAEEACIRTSHLKQAAPOOQGLRAOSQPOQVQKIHRE 991
Db 932 SOSTN-----HIYNSLAN-----CSPNRPVVDHVS 959
RESULT 9
O97744 PRELIMINARY; PRT; 964 AA.
AC O97744;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Mast/stem cell growth factor receptor (Fragment).
GN KITI*0101.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OC NCBI_TaxID=9823;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=HAMPSHIRE BREED; TISSUE=LEUKOCYTE;
RX MEDLINE=98391767; PubMed=9724328;
RA Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,
RA Moller M., Edfors-Lilja I., Andersson L.;
RT "Molecular evolution of the dominant white phenotype in the domestic
RL pig.";
RL Genome Res. 8:826-833(1998).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=HAMPSHIRE BREED; TISSUE=LEUKOCYTE;
RA Marklund S.;
RL Thesis (1997), Department of Animal Breeding and Genetics,
RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; AJ223228; CAA11196.1; -;
```

```
DR HSP; P11362; IFGK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_pkinase; 2.
DR SMART; SM00409; Ig; 1.
DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR Glycoprotein; Immunoglobulin domain; Phosphorylation; Receptor;
KW Transmembrane; Tyrosine-protein kinase.
FT NON_TER 964 964
SQ SEQUENCE 964 AA; 108287 MW; 90E72EDFAB1358B7 CRC64;

Query Match 22.9%; Score 1208; DB 6; Length 964;
Best Local Similarity 32.6%; Pred. No. 2.8e-97;
Matches 319; Conservative 168; Mismatches 362; Indels 130; Gaps 32;

QY 38 LISHEHNGSAGPSSRYMVRGSPEDLOCTPRQSEGTVEATVEAESGSLTQVOLA 97
Db 16 LLLRVOTGSS--QPSV-----SPEEL-----SPSIHPAKSELIVSAGD-EIRLFCT 59
QY 98 TPGDLSCLVAFKHSGLGCPHFEDLNQNGIVSMALNTVETQAGYLLHIOISERA----- 151
Db 60 DPG--SVKWTFE--TLG-----QLSENTHAEIV-----EKAEMNTG 93
QY 152 NYTVL-----FTVNVRTQ--LYVLRREYFRKMNQDAL-----LCISE 188
Db 94 NYTCNTEGGLSSIIYFVRDPEKLFVDPPLYGK-EDNDALVRCPLTDPEVTNYSLTGCE 152
QY 189 GVPEP-TVVEWLCSSHRESCKEKGPAVVRKEEVHLFEGTDIRC--CARNALGRE-CTK 244
Db 153 GKPLPKDLTEV-----ADPKAGITIRNVKREYHRL-----CLHCSANQOGKSVLSK 198
QY 245 LETIDLNQAPQS-----TLPQLFLKVGPELWIRCKAIHVNHGFLTWELDKALEEGSY 298
Db 199 KFTLKVRRAIRAVPVAVASVYLLRREGEEFAVMCLIKDVSSVSDSMW--IRENSQTKAQV 257
QY 299 FEMSTYSTNRTMIRILLAFVSSVGRNDTGYTCSKSHPSQSALVTILE---KGFNA-T 354
Db 258 KRNSWHOGDFNLRQERLTISARVNDSGVFMCVANNFTGSGANVTTLTVVVDKGFNIFF 317
QY 355 SSQEEYEDIDYKFCFSVRFKAYPR-IRCTWIFSQ--ASFPCEORGLGDSISKFCDHK 411
Db 318 MNMTTFVNDGEDVDLIVEYEAYPEKHROWIYMNRTATDKWEDYPKSENEISRIYSEL 377
QY 412 N-----KPEXIFYAENDDAQFTKMTLNIRKKPOVLAN--ASASQSCSSDGYPLPS 462
Db 378 HLTRLKTEGTYTFLVSNADVNSVTFNVYNTKPELLTHDRLMNGMLQCAVAGFPPEPT 437
QY 463 WTWKCKSDKSPNCTEEI-PEGVWNNKANRKFVGOWVSSSTLNMSEAGKGLLVKCCAYNSM 521
Db 438 IDWYFCPTQRCVSPVPGVDVQIQNSVSPFGKLVIHSSIDYSAFKHNGTVECRAYNDV 497
QY 522 GTSCTETFLNSPGPFPIQDNISFYATIGLCPLPIV-----VLIVLICHYKKQFRYES 575
Db 498 GKS--SAFFN---FAFKEQIHAHTLTPLTLLIGFVIAAGMCCIIMVILTYLQKPMYEV 551
QY 576 QLQMI-QVTGPLDNEYFYVDFRDYEDLKWEEFRENLEFGKVLGSGAFGRVYNNATAYGIS 634
Db 552 QWKVVEEING---NNYVIIDPTQLPYDHKWEFFPNRLSFGKLTGAGAFKVVETAYGLI 608
QY 635 KTGSIOVAVKMLKEKADSCKEALMSKLMTHLGHHDNIVNLLGACTLSGPPVYLIFEY 694
```

Db 609 KSDAAMTVAVKMLKPSAHLTEREALMSELKVLVSLGNHNMNIVNLLGACTIGGPTLVITEY 668  
 QY 695 CCYGDLLNLYLRKRREKFRHTWTETFEKHNFSYPTFOAHSNSSMPGSRVQLHPPLDQLS 754  
 Db 669 CCYGDLLNLYLRKRREKFRHTWTETFEKHNFSYPTFOAHSNSSMPGSRVQLHPPLDQLS 728  
 QY 755 GNGNSIHSDEIEYENOKRLAEEDLNVLTFEDLLCFAYQVAKGMFEFLFKSCVHRD 814  
 Db 729 PTKADKRRSARIGSYIERDVTPIAIMEDELALDLEDLLSFSYQVAKGMFLASKNCIHRD 788  
 QY 815 LAARNVLTHGVKVICDFGLARDILSDSSYVVRGNARLPVKWMAPESLFEGIYIKSDV 874  
 Db 789 LAARNILTHGRITKICDFGLARDIKNSNYVVKGNARLPVKWMAPESLFNCVYTFESDV 848  
 QY 875 WSYGILLWEIFSLGVNPPGIPVDANFYKLIQSGFKMEOPFYATEGIFVMOCSWAFDSR 934  
 Db 849 WSYGIFLWELFSLGSSPPGPNVDSKFKYMKIEGFRMLSPHAPAEWYDIMTKWDADPL 908  
 QY 935 KRSPSPNLTSLGCOLAFA 953  
 Db 909 KRPTFKQIVQLIEKOISES 927

RESULT 10  
 Q9TQ01 PRELIMINARY; PRT; 964 AA.  
 AC Q9TQ01;  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Mast/stem cell growth factor receptor (Fragment).  
 GN K11L\*0201.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RA STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;  
 RA Marklund S.;  
 RL Thesis (1997), Department of Animal Breeding and Genetics,  
 RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.  
 [2]  
 RC SEQUENCE FROM N.A.  
 RA STRAIN=YORKSHIRE / LANDRACE CROSS; TISSUE=LEUKOCYTE;  
 RX MEDLINE=98391767; PubMed=9724328;  
 RA Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,  
 RA Moller M., Edfors-Lilja I., Andersson L.;  
 RT "Molecular evolution of the dominant white phenotype in the domestic  
 RT pig";  
 RL Genome Res. 8:826-833(1998).  
 CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -|- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC EMBL; AJ223229; CAAL1197.1; .  
 DR HSSP; P11362; IFGK.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR003599; Ig.  
 DR InterPro: IPR003600; Ig\_like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR001824; RTKkinaseIII.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00047; Ig; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Euk\_pkinase; 2.  
 DR SMART; SM00409; IG; 1.  
 DR SMART; SM00410; IG\_like; 1.  
 DR SMART; SM00219; TyrcK; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 2.  
 DR ProSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR ProSITE; PS00107; PROTEIN\_KINASE\_DOM; 1.  
 DR ProSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR ProSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.

KW Glycoprotein; Immunoglobulin domain; Phosphorylation; Receptor;  
 KW Transmembrane; Tyrosine-protein kinase.  
 FT NON\_TER 964  
 SQ SEQUENCE 964 AA; 108315 MW; 9963046201358A8 CRC64;  
 Query Match 22.9%; Score 1207; DB 6; Length 964;  
 Best Local Similarity 32.6%; Pred. No. 3.5e-97;  
 Matches 319; Conservative 168; Mismatches 362; Indels 130; Gaps 32;  
 QY 38 LLSHENNGSSAKPSSYRMVGRGSPEDLOCTPRRQSEGVYEAATVEVAESGISTLQVOLA 97  
 Db 16 LLRLVQTGSS--QPSV-----SPEEL-----SPSHPAKSELIVSAGD-EIRLFT 59  
 QY 98 TPGDLSCLWVFKHSSLCGCPHFEDLQNRGIVSMAILNVETQAGEYLLHIOSERA----- 151  
 Db 60 DPG--SVKWTFE--TLG-----OLSENTHAEWIV-----EKAEAMNTG 93  
 QY 152 NYTVL-----FTVNVRTDQ-LYVLRPRPYFRKMNODAL-----LCISE 188  
 Db 94 NYTCTNEGGLSSIIYFVRDPEKFLVLDPLYGK-EDNALVRCPLTDPVTNYSLTGCE 152  
 QY 189 GVPEP-TVEMVLCSSHRESCKEKGPAVVRKEKVLHELFCTDIRC--CARNALGRE-CTK 244  
 Db 153 GKPLPKDLTFV-----ADPKAGITIRNVKREYHRL-----CLHSANOGSKVLSK 198  
 QY 245 LETIDLNOAPQS-----TLPOLFLKVGPELWIRCAIHVHNGFGTLWELEDKALEGSY 298  
 Db 199 KFTLVRAAIRAVPVVAVKASYLLREGEFEFVAVMCLIKDVSSVDSMW-IRENSQTKAQV 257  
 QY 299 FEMSTYSTNRTMIRILLAPVSSVGRNDTCYTCSSSKHPQSOSALVILE---KGFINA-T 354  
 Db 258 KNSWHQGDNFLEORLTISSARVNDSGVFMCIYANTFGSANVTTLLEVVDGFGFINIFP 317  
 QY 355 SSQEEYIDPEKFCFSVRKAYPR-IRCTWIFSQ--ASFPCERQGLEDCYISKFCDHK 411  
 Db 318 MNNTTVFVNDGEDVDLIVEYEAYPKEPHQWIYMNRTATDKWEDYKPKSENIYRIVSEL 377  
 QY 412 N-----KPEYIFYAENDDAQFTKMTFLNIRKPKQVLAN--ASASQSCSSDGYPLPS 462  
 Db 378 HLRLKGTGGTYTFLVSNADYNSVTFENVYNTKPEILTDLRLMGLMGLQCVAAAGPEPT 437  
 QY 463 WTWKCKDSKSPNCTEEI-PEGVYNNKANKRVFGVOMVSSSTLANMSEAGGLLVKCCAYNSM 521  
 Db 438 IDWYFCPTGTEQRCVPVGVVDVQIIONSSVSPFGKLVHSSIDIYSAFKHNGTVECRAYNDV 497  
 QY 522 GTSCETIFLNSPGPPFFIQDNISFYATIGLCPLPFIV-----VLIVLCHKYKKQFRYES 575  
 Db 498 GKS--SAFFN---FAKQEQIHAHTLFTPLLLIGFVIAAGMCIIVMLTYKYLQKPMYEV 551  
 QY 576 QLOMI-QVTGPDLDNIFYVDFRDYEDLKWEPPEMLNLEFGKVLGSGAFGRVMNATAYGIS 634  
 Db 552 QKVVVEEING---NNYVYIDPTQLPDYDHKWEPRNRLSEFGKTLGAGAFGVVEATAYGLI 608  
 QY 635 KTVGSIQVAVKMLKEKADSCKEALMSELKMTHTLGHHDNIYNLLGACTLSGPGVYLIFEY 694  
 Db 609 KSDAAMTVAVKMLKPSAHLTEREALMSELKVLVSLGNHNMNIVNLLGACTIGGPTLVITEY 668  
 QY 695 CCYGDLLNLYLRKRREKFRHTWTETFEKHNFSYPTFOAHSNSSMPGSRVQLHPPLDQLS 754  
 Db 669 CCYGDLLNLYLRKRREKFRHTWTETFEKHNFSYPTFOAHSNSSMPGSRVQLHPPLDQLS 728  
 QY 755 GNGNSIHSDEIEYENOKRLAEEDLNVLTFEDLLCFAYQVAKGMFEFLFKSCVHRD 814  
 Db 729 PTKADKRRSARIGSYIERDVTPIAIMEDELALDLEDLLSFSYQVAKGMFLASKNCIHRD 788  
 QY 815 LAARNVLTHGVKVICDFGLARDILSDSSYVVRGNARLPVKWMAPESLFEGIYIKSDV 874  
 Db 789 LAARNILTHGRITKICDFGLARDIKNSNYVVKGNARLPVKWMAPESLFNCVYTFESDV 848  
 QY 875 WSYGILLWEIFSLGVNPPGIPVDANFYKLIQSGFKMEOPFYATEGIFVMOCSWAFDSR 934  
 Db 849 WSYGIFLWELFSLGSSPPGPNVDSKFKYMKIEGFRMLSPHAPAEWYDIMTKWDADPL 908



```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=CEREBELLUM;
RA  London C.A., Galli S.J., Yuuki T., Hu Z.-Q., Helfand S.C.,
RA  Gaissler E.N.;
RT  "Spontaneous canine mast cell tumors express tandem duplications in
RT  the proto-oncogene c-kit.";
RL  Exp. Hematol. 0:0-0(1999).
CC  -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC  PROTEIN KINASES.
DR  EMBL; AF099030; AAD28369.1; -.
DR  HSP; P11362; IFGK.
DR  InterPro; IPR000719; Euk_pkinase.
DR  InterPro; IPR003599; Ig.
DR  InterPro; IPR003600; Ig-like.
DR  InterPro; IPR003006; Ig_MHC.
DR  InterPro; IPR001824; RTKinaseIII.
DR  InterPro; IPR001245; Tyr_pkinase.
DR  Pfam; PF00047; Ig; 2.
DR  Pfam; PF00069; pkinase; 1.
DR  ProDom; PD000001; Euk_pkinase; 2.
DR  SMART; SM00409; IG; 2.
DR  SMART; SM00410; IG-like; 1.
DR  SMART; SM00219; TyKc; 1.
DR  PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
DR  PROSITE; PS00011; PROTEIN KINASE_DOM; 1.
DR  PROSITE; PS00109; PROTEIN KINASE_TYR; 1.
DR  PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR  ATP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
KW  Receptor; Transferase; Transmembrane; Tyrosine-protein kinase.
SQ  SEQUENCE 978 AA; 109651 MW; 5AC31E5AC4E9910F CRC64;

Query Match      22.9%; Score 1205; DB 6; Length 978;
Best Local Similarity 31.5%; Pred. No. 5.4e-97;
Matches 329; Conservative 167; Mismatches 387; Indels 160; Gaps 33;

QY  14  LVLVLSVILETVNQDLPVVKLVLSHENNGSSAGRPSSYRMVGRSPEDLQCTPRQSE 73
DB  11  LCVLLLLLVQVQSGQPSV-----SPGEPFL----- 38

QY  74  GTVEAATVEAEGSGITLQVQLATPGDLSCWLVFKSISSGCPHFQDLQNRGIVSMATLN 133
DB  39  PSIHAKSELIVSGD-ELRLSCTDPGVK-WTFE--TLG-QLNENTHNEWIEKA--- 89

QY  134  VTQAGEYLLHQSERANTVLTFTVNRD-TQLYVLRPRYFRKMNQDALLCISEGVPE 192
DB  90  ----EAGHTGNYTCTNRDGLRSIYFVROPAPKLFVLDPLYGKGNLTVRC----- 138

QY  193  PTVEWVLCSSHRESCKEGPAVVRKEEVLHELFGTDIR-----C--CARNALGRE 241
DB  139  PLTDPEVTNYSLRG--EGKPLPKDLFTFADPKAGITIRNVKRYHRLCLHCSADQKRT 196

QY  242  -CTKLFTIDLNQAPQS-----TLPQLFLKVGPELWIRCKRAIHVNHGFLTWLEDKALE 294
DB  197  VLSKKFTLKRAIRAVPVVSVKTSLLKEGEAFSVMCFIKDVSSFVDSMW-----IK 250

QY  295  EGSYFEMSTYST-----NRTMIRILLAFVSGVRNDTGYYTCCSSKHPSQSALVTILE-- 347
DB  251  ENSOTNAQTQSNWSHHGDFNERQEKLIISARVNDSGVEMCYANNTFGSANVTITLEV 310

QY  348  -KGFINA-TSQSEYEIDPDYKCFSVRFKAYPR-IRCTWIFSQASF--PCEQRGLEDGY 402
DB  311  DKGFINIFPMWMTTIFVNDGENVDLIYVEYAYPKPEQOQIYMNRTETDKWEDYPKSDNE 370

QY  403  SISKFCDH-----KNKPGEIFYAEANDAQTFTKMTLNLIRKKPOVLANASQA--SC 453
DB  371  SNRYVSELHLRLKNGEGGTYTFQVNSDVNSVTVTNVYNTKPEILTHESLTNGMLQC 430

QY  454  SSDGYPPLPSWTWKKCDKSPNCTBEI-PEGVWKNKARKKVFQGWSSSTLNMSAGKGLL 512
DB  431  VVAGFPPEPAVDWFCPCAEQRCQSVIPGMDVQMGNSLSFGKLVQSSIDYSAFKHNGT 490

```

```

QY  513  VKCCAYNSMGTSCETIFLNSPGPFPIQDN-----ISFYATIGLCLPFIIV 558
DB  491  VECRAYNNVGRS--SAFTN-----FAFKGNSKEQIHPHTLTPTLLIGFVIAAGM---MCI 540

QY  559  LIVLICHYKKOFRYESQLQMI-QVTGPDLDNEYFYDFRDYEDYDLKWEFPRENLEFGKVL 617
DB  541  IVMLITYKYLQKPMYEQWKVVEENG---NNVYVIDPTQLPYDHWKFPNRLSFGKTL 597

QY  618  GSGAFGRVWNTAYIGISKTVGSIQVAVKMLKEKADSCKEALMSELKMMTHLGHNDINIV 677
DB  598  GAGAFGKVEATAYGLIKSDAAAMTAVAKMLKPSAHLTEREALMSELKVLSYLGNHNNIV 657

QY  678  LLGACTLSGPVYLIFECCYGDLLANVLRSKREKFEHTWTETFEKHNFFSSYPYTFQAHNSSS 737
DB  658  LGACTVGGPTLVTITEYCCYGDLLNLFRRKRDSFCISKQEDHGEVALYKNLLHSKSSCS 717

QY  738  MPGSREVOLHPPDLQDLSGFNGNSIHSEDEIEYENKRLAEEDN-----VLTF 788
DB  718  DSTNEYMDMKP-----GDSYVVPVKADKRRSRIGSYIERDVTYPAIMDEDELALDL 768

QY  789  EDLLCFAYQVAKMEFEFKSCVHRDLAARNVLVTHGKVKVYKICDFGLARDILSDSSVYVR 848
DB  769  EDLLSFSYQVAKMAFLASKNCHRDLAARNILLTHGRITKICDFGLARDIKNDSNVYVK 828

QY  849  GNARLPVKWAPESLFEGLIYTIKSDVWSYGIILLWEIFSLGVNRPYGPVVDANFYKLIQSC 908
DB  829  GNARLPVKWAPESIFNCVYTFESDVWSYGIIFLWELFSLGSSPYGPMGVDSFYKMIKEG 888

QY  909  FKMEOPFATEGIYFVMQSMFAFDSKRPSFNLTSLFCQLAAEAEACIRTSIHLPKQA 968
DB  889  FRLMSPEHAPAEYIMYKTCWDADPLKRTSKQIVQLIEKQISDSTN-----HTYSNL 941

QY  969  APQQRGGLRAOSPORQVKIHRER 991
DB  942  AN-----CSPNPERPVVDHSVR 958

RESULT 13
Q98S02
ID  Q98S02      PRELIMINARY;      PRT;      977 AA.
AC  Q98S02;
DT  01-JUN-2001 (TReMBLrel. 17, Created)
DT  01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT  01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE  Receptor tyrosine kinase Fms.
OS  Danio rerio (Kerr's danio).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC  Cyprinidae; Danio.
OX  NCBI_TaxID=38750;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=21359118; PubMed=11466528;
RA  Parichy D.M., Johnson S.L.;
RT  "Zebrafish hybrids suggest genetic mechanisms for pigment pattern
RT  diversification in Danio.";
RL  Dev. Genes Evol. 211:319-328(2001).
CC  -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC  PROTEIN KINASES.
DR  EMBL; AF324480; AAKL5302.1; -.
DR  HSP; P11362; IFGK.
DR  InterPro; IPR000719; Euk_pkinase.
DR  InterPro; IPR003599; Ig.
DR  InterPro; IPR003598; Ig_C2.
DR  InterPro; IPR003006; Ig_MHC.
DR  InterPro; IPR001824; RTKinaseIII.
DR  InterPro; IPR002290; Ser_thr_pkinase.
DR  Pfam; PF00047; Ig; 4.
DR  Pfam; PF00069; pkinase; 1.
DR  ProDom; PD000001; Euk_pkinase; 2.
DR  SMART; SM00409; IG; 4.

```

```
DR SMART: SM00408; IGC2; 1.
DR SMART: SM00220; S_TKC; 1.
DR SMART: SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;
KW Phosphorylation; Receptor; Transferase; Transmembrane.
FT VARIANT 31 P -> L.
FT VARIANT 148 L -> F.
FT VARIANT 218 R -> K.
FT VARIANT 263 I -> M.
FT VARIANT 490 R -> K.
FT VARIANT 594 E -> A.
FT VARIANT 614 G -> R.
FT VARIANT 926 N -> T.
FT VARIANT 934 A -> T.
FT VARIANT 938 A -> V.
SQ SEQUENCE 977 AA; 110294 MW; AAB2FF2A20D66090 CRC64;

Query Match 22.9%; Score 1204.5; DB 13; Length 977;
Best Local Similarity 32.2%; Pred. No. 5.9e-97;
Matches 334; Conservative 166; Mismatches 363; Indels 173; Gaps 36;

Qy 12 LLLVLLVSMILETVTNODLPVVKLVSHENNGSSACKPSSYRMVRCGSPEDLOCTPRRQ 71
Db 12 LLLVLLVSMILETVTNODLPVVKLVSHENNGSSACKPSSYRMVRCGSPEDLOCTPRRQ 71
Qy 5 LLLLIGI---LFGQVGWSEFOIRL-----NSGAPAG--TDVILDGSPLOLVC----- 48
Db 5 LLLLIGI---LFGQVGWSEFOIRL-----NSGAPAG--TDVILDGSPLOLVC----- 48
Qy 72 SEGTVEAATVEAEGSITITQVQLATPGDLSCLVFVKHSSIGCOPHFQDLQNRGIVSMAI 131
Db 72 SEGTVEAATVEAEGSITITQVQLATPGDLSCLVFVKHSSIGCOPHFQDLQNRGIVSMAI 131
Qy 49 -EG-----DGPVTLPLRA-----KH-----KRYISKEVGKIRSPRV 79
Db 49 -EG-----DGPVTLPLRA-----KH-----KRYISKEVGKIRSPRV 79
Qy 132 LNVETQAGEY-LHIOISERANYVLFVNVNVDTO-LVY---LRPRYFKMENQDALL-C 185
Db 132 LNVETQAGEY-LHIOISERANYVLFVNVNVDTO-LVY---LRPRYFKMENQDALL-C 185
Qy 80 EKATVDFGTGKCVKVINENSLSSVHFVRDSRVLFVSPSITLRYVRK-EGEDLLLP 138
Db 80 EKATVDFGTGKCVKVINENSLSSVHFVRDSRVLFVSPSITLRYVRK-EGEDLLLP 138
Qy 186 ISEGVPEPTVENVLCSSHRESCKEKG-PANVVRKEKVL---HELFGTDIRCCARNALGR 240
Db 186 ISEGVPEPTVENVLCSSHRESCKEKG-PANVVRKEKVL---HELFGTDIRCCARNALGR 240
Qy 139 L---LTDPDATDFMRDMONGSAAPYGMNATEDPRKGVLRNVHPGNADYICSAIRGAE 195
Db 139 L---LTDPDATDFMRDMONGSAAPYGMNATEDPRKGVLRNVHPGNADYICSAIRGAE 195
Qy 241 ECTKLFTIDLNAQPOSTLPQLFLK-----VGEPLWIRCKAIHNVHGFGLFWELEDKAL 293
Db 241 ECTKLFTIDLNAQPOSTLPQLFLK-----VGEPLWIRCKAIHNVHGFGLFWELEDKAL 293
Qy 196 KVSIFSIINIQLRFP-PYVLLRNEVYKLVGERLQISCTNNFNFFYVNTWTHSSRRL 254
Db 196 KVSIFSIINIQLRFP-PYVLLRNEVYKLVGERLQISCTNNFNFFYVNTWTHSSRRL 254
Qy 294 EGSYFEMSTYSTNTMIRILLAFVSSVGRNDTGYTCCSSKHPSQSALVT---ILEKGF 350
Db 294 EGSYFEMSTYSTNTMIRILLAFVSSVGRNDTGYTCCSSKHPSQSALVT---ILEKGF 350
Qy 255 PKAE--EKSTIEGDRLAIESILT-IPSVQOSDNGNITCTGQNEAGANSSTTQLLYVDEPY 311
Db 255 PKAE--EKSTIEGDRLAIESILT-IPSVQOSDNGNITCTGQNEAGANSSTTQLLYVDEPY 311
Qy 351 INAT---SSQEEY-----EIDPYEKFCFSVRKPYRIRC-----TWIFSQASFPCQR-- 396
Db 351 INAT---SSQEEY-----EIDPYEKFCFSVRKPYRIRC-----TWIFSQASFPCQR-- 396
Qy 312 IRLSPKLSKLTTHRLSTIEVSEGDVDGLVLEATPPLTSHKWEPTTSINASLP-ENRY 370
Db 312 IRLSPKLSKLTTHRLSTIEVSEGDVDGLVLEATPPLTSHKWEPTTSINASLP-ENRY 370
Qy 397 GLEDGYISKFCDHKN--KPGEYIFEAENDDAQFTKMTFLNIRKKPOVLAN-ASASQASC 453
Db 397 GLEDGYISKFCDHKN--KPGEYIFEAENDDAQFTKMTFLNIRKKPOVLAN-ASASQASC 453
Qy 371 NHNDYEALLFLKRLNFEIEIGYTLNVNKSMSASITFDIKMYTPVARVKWENVTLS 430
Db 371 NHNDYEALLFLKRLNFEIEIGYTLNVNKSMSASITFDIKMYTPVARVKWENVTLS 430
Qy 454 SSDGYPLPSWTWKCKSDKSPNCTEI----PEGVNNKANRVKVGQWSSSTLNNSKAG 509
Db 454 SSDGYPLPSWTWKCKSDKSPNCTEI----PEGVNNKANRVKVGQWSSSTLNNSKAG 509
Qy 431 RSYGYPAPSILWYQCTGIRTTCPENTDLPQTOTQTOVEPKESFGAVGESVLTGV-PNR 489
Db 431 RSYGYPAPSILWYQCTGIRTTCPENTDLPQTOTQTOVEPKESFGAVGESVLTGV-PNR 489
Qy 510 GLLVKCCAYNSMGTSCTEIFLNSPGFPFIQDNISFYATIGLCLPFIV--VLIVLIC 567
Db 510 GLLVKCCAYNSMGTSCTEIFLNSPGFPFIQDNISFYATIGLCLPFIV--VLIVLIC 567
Qy 490 RMTVVVCAFNVLVGGSDFTSMD-----VSDQIFTSAMCGSTVAMVGLLLIIMYIKY 542
Db 490 RMTVVVCAFNVLVGGSDFTSMD-----VSDQIFTSAMCGSTVAMVGLLLIIMYIKY 542
Qy 568 KQPFYEQSOLQMIQVTPGLDNIEYFVDFRDYEDLKWPEPPRENLEFGKVLGSGAFGRV 627
Db 568 KQPFYEQSOLQMIQVTPGLDNIEYFVDFRDYEDLKWPEPPRENLEFGKVLGSGAFGRV 627
Qy 543 KQPFYEQSOLQMIQVTPGLDNIEYFVDFRDYEDLKWPEPPRENLEFGKVLGSGAFGRV 600
Db 543 KQPFYEQSOLQMIQVTPGLDNIEYFVDFRDYEDLKWPEPPRENLEFGKVLGSGAFGRV 600
Qy 628 ATAYGISKTGVSIOYAVKMLKEKADSCKEALMSKLMTHLGHHDNIYNLLGACTLSGP 687
Db 628 ATAYGISKTGVSIOYAVKMLKEKADSCKEALMSKLMTHLGHHDNIYNLLGACTLSGP 687
Qy 601 ATAYGLGKREDNITGVAVKMLKASAPDEREALMSKILSLHGLQHKNIYNLLGACTHGGP 660
Db 601 ATAYGLGKREDNITGVAVKMLKASAPDEREALMSKILSLHGLQHKNIYNLLGACTHGGP 660
```

```
Qy 688 VYLIPEYCYGDLNLVLRSKREKF-----HRTW-----TEI 718
Db 688 VYLIPEYCYGDLNLVLRSKREKF-----HRTW-----TEI 718
Qy 661 VLVITEYCHGDLNLFRLSKAENFLNFVMTIPNFPEPMTDYKNVSTERMFVRSDSGISST 720
Db 661 VLVITEYCHGDLNLFRLSKAENFLNFVMTIPNFPEPMTDYKNVSTERMFVRSDSGISST 720
Qy 719 FKEHNFSSYPTQAHNSNSMPSGSRVQLHPPLDQLSGFNGNSIHSEDEIEYENQKRLAE 778
Db 719 FKEHNFSSYPTQAHNSNSMPSGSRVQLHPPLDQLSGFNGNSIHSEDEIEYENQKRLAE 778
Qy 721 CSDHYLDMPVTSRPTNSTLSDSSEQ----- 747
Db 721 CSDHYLDMPVTSRPTNSTLSDSSEQ----- 747
Qy 779 BEEDNLVLTFFDLICFAYOVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARD 838
Db 779 BEEDNLVLTFFDLICFAYOVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARD 838
Qy 748 --EDSNPLDMDDLLRFSQVAGLDFLAANKCIHRDMAARNVLLTNSRVAKICDFGLARD 805
Db 748 --EDSNPLDMDDLLRFSQVAGLDFLAANKCIHRDMAARNVLLTNSRVAKICDFGLARD 805
Qy 839 ILSDSYVYVRGNARLPVKWMAPESEFEGYITIKSDVMSYGILLWEIFSLGVNPIPGIPVD 898
Db 839 ILSDSYVYVRGNARLPVKWMAPESEFEGYITIKSDVMSYGILLWEIFSLGVNPIPGIPVD 898
Qy 806 INWDSYVYVRGNARLPVKWMAPESEFEGYITIKSDVMSYGILLWEIFSLGVNPIPGIPVD 865
Db 806 INWDSYVYVRGNARLPVKWMAPESEFEGYITIKSDVMSYGILLWEIFSLGVNPIPGIPVD 865
Qy 899 ANFYKLIQSGFKMEOPFYATEGIYFVWQSCWAFDSKRKSPNLTSLFGCOLAE-AEEAC 957
Db 899 ANFYKLIQSGFKMEOPFYATEGIYFVWQSCWAFDSKRKSPNLTSLFGCOLAE-AEEAC 957
Qy 866 SKFYKMIKGYQMSRPDPFAPPENYTIMKMCWNLDAAERTFSKISQIMQIRMLGETSEQOD 925
Db 866 SKFYKMIKGYQMSRPDPFAPPENYTIMKMCWNLDAAERTFSKISQIMQIRMLGETSEQOD 925
Qy 958 IRTSHLPKQAAPQOR 973
Db 958 IRTSHLPKQAAPQOR 973
Qy 926 NOEYKNIPAEAEQO 941
Db 926 NOEYKNIPAEAEQO 941

RESULT 14
Q9TTD7 PRELIMINARY; PRT; 948 AA.
AC Q9TTD7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Stem cell factor receptor (Fragment).
GN C-KIT.
OS Trichosurus vulpecula (Brush-tailed possum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
OX NCBI_TaxID=9337;
RN [1]
RP SEQUENCE FROM N.A.
RA Lawrence S.B., Greenwood P., Tisdall D.J., McNatty K.P., Fidler A.E.;
RT "Partial cDNA sequence of the c-kit homologue of brushtail possum
RT (Trichosurus vulpecula).";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF131209; AAF22141.1;
DR HSP; P11362; IFGK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; ig; 3
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR Immunoglobulin domain; Receptor; Tyrosine-protein kinase.
KW NON_TER 948
SQ SEQUENCE 948 AA; 106323 MW; 708E0258FFC4D07D CRC64;

Query Match 22.8%; Score 1201; DB 6; Length 948;
Best Local Similarity 33.0%; Pred. No. 1.2e-96;
Matches 314; Conservative 148; Mismatches 356; Indels 134; Gaps 29;

Qy 77 YEATVEAEGSITITQVQLATPGDLSCLVFVKHSSIGCOPHFQDLQNRGIVSMA 130
Db 77 YEATVEAEGSITITQVQLATPGDLSCLVFVKHSSIGCOPHFQDLQNRGIVSMA 130
Qy 33 YSSPSITPARS---QLTVNFGSEIKUSCDLHFNVTFFENAL----HLE-SSRTTTL 84
Db 33 YSSPSITPARS---QLTVNFGSEIKUSCDLHFNVTFFENAL----HLE-SSRTTTL 84
Qy 131 ILNVETQAGEYLLHIOISERANYVLFVNVNVDTO-LVY---LRPRYFKMENQDALL- 183
Db 131 ILNVETQAGEYLLHIOISERANYVLFVNVNVDTO-LVY---LRPRYFKMENQDALL- 183
```

Db 85 TNAKAODTCRYTC---TNKGLSSSIYFVKDPKILFLDLRLLYNEGSDALVNCPTVD 141  
QY 184 -----LCISEGVPEP-----TVEWVLCSSHRESCKEKGPAVVRKEEVLH 223  
Db 142 PEVNTFLRLCDGKPLPKDLTLPIQKGTIRNKV-RSHNKIC----- 184  
QY 224 ELFGTDIRCCARNALGRECTKLTIDLNOAPOSTLPOL-----FLKVGEPWLIRCKAI 276  
Db 185 -----FOCSAYODGOLKLSDRMTLKVRAPOS-VPEVSLIQTNYLLREGETFOATCMIK 237  
QY 277 HVNHGFLTWELEDKALEEGSYEMSTYSNRTMIRILLAPFVSVGRNDGTGYTCSSSK- 335  
Db 238 DVASSVVMKIKNNRT---STHTQSHSGDYAYERODILITISPRVNDSGVTCFANNT 294  
QY 336 --HPSOSALVTILEKGFINA-TSQEYEYEDPYEKFCSVRFKAYPR-IRCTWIFSOASF 391  
Db 295 FGANVTATLKVVEKGFINFPKMNNTIFINDGENIDLVVEYEAFFKPEHLQWIMNGTV 354  
QY 392 PCORGLEDGYS-----ISKCDHNK-----PGEYIFYAENDDAQFTKMFLLNI 436  
Db 355 TDKW---DDYTRPGSESTIRYISEL--HLNRLKGTGGPYTFVSNSDSTSVTFKVVY 408  
QY 437 RKPQVLNANASOA--SCSDGYPLSTWTKKCDKSPNC-TTEIPEGVMKKANKRVF 493  
Db 409 KTRPEILTSRLMNLQCVAAAGPEPTIDWYFCPGTEQRCSTSVLPMDVKNTNSMLPPE 468  
QY 494 GOMVSSSTLANKSEAGKLLVKKCAYNMGTSCETIFLNSGPPFFIODNISFYATIGLCL 553  
Db 469 GKIVESTIDSSAFRYNGTVECKASNDVGKS--SAFTN----FAIKQITSHLTFTPLLI 522  
QY 554 PFIV-----VLIVLICHYKKOFRYESQLOMI-QVTPGLDNEYFYVDFRDEYEDLWKEF 606  
Db 523 GFVVAAGLMCVIILTYKYFQRPMEYQVQVVEING---NNVYIDPTQLPDYDHWKEF 579  
QY 607 PRENLFGKVLGSGAFGRVWNAATYISKTGVSIOQAVKMLKADKADCEKALMSLKMM 666  
Db 580 PRNRLSFGKTLGAGAFGVVEATAYGLFKSDAAMTAVAKMLKPSAHLTEREALMSLKVL 639  
QY 667 THLGHDNIIVNLGACLTSLGPVLIIFYCCYGDLLNLVLRKREKFRHTWEIFKEH---- 722  
Db 640 SYLGNHNIIVNLGACLTSLGPVLIIFYCCYGDLLNLVLRKREKFRHTWEIFKEH---- 722  
QY 723 NFSYPTFOAHNSMPGSRV-VOLHPPLDQLSGFNGNSHSEDEIEYENOKRLAEEEEE 781  
Db 694 EALYKNNLQKESSCDGANEDYMDKPGVYVPTKREKRSARVGSYIERDVTYATMED 753  
QY 782 DLNLTFFEDLLCFAYQVAKMELEFKSVHRDLAARNVLVTHGKVYKICDFGLARDILS 841  
Db 754 DELALDIEDLLSPSYQVAKGMSFLASKNCIHRDLAARNILLTHGRITKICDFGLARHIK 813  
QY 842 DSSVYVGNARLPVKWAPESLFEIGYITIKSDVWSYGILLWEIFSLGVNYPGPVDPANF 901  
Db 814 DSNVYVGNARLPVKWAPESIFNCVTFESVWSYGIWELFSLGSSYPGMPVDSKF 873  
QY 902 YKLTQSGFKMEQPFYATEGIFYFVQSCWAFDSRRKPSFNLTSFLGCOLAEA 953  
Db 874 YKMKEGFRLMSPECAPPENYEMKSCWNEPDLQRPFTFKLIVLIEQOQLDS 925  
  
RESULT 15  
Q98SU4  
ID Q98SU4 PRELIMINARY; PRT; 977 AA.  
AC Q98SU4;  
DT 01-JUN-2001 (TReMBLrel. 17, Created)  
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE Receptor tyrosine kinase Fms  
OS Danio albolineatus (pearl danio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=27699;  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=21359118; PubMed=11466528;  
RA Parichy D.M., Johnson S.L.;  
RT "Zebrafish hybrids suggest genetic mechanisms for pigment pattern  
diversification in Danio.";  
RL Dev. Genes Evol. 211:319-328(2001).  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
PROTEIN KINASES.  
DR EMBL; AF324478; AAKI5300.1; -.  
DR HSSP; P11362; IFGK.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR003600; Ig\_like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR001824; RTkinaseIII.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR Pfam; PF00047; Ig\_4.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 2.  
DR SMART; SM00409; IG; 4.  
DR SMART; SM00408; IGC2; 1.  
DR SMART; SM00410; IG\_Like; 1.  
DR SMART; SM00220; S\_TKc; 1.  
DR SMART; SM00219; TyrcK; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
KW Phosphorylation; Receptor; Transference; Transmembrane.  
FT VARIANT 57 57 L -> V.  
FT VARIANT 164 164 A -> V.  
FT VARIANT 325 325 R -> P.  
FT VARIANT 326 326 G -> C.  
FT VARIANT 327 327 L -> P.  
FT VARIANT 328 328 S -> A.  
FT VARIANT 328 328 S -> F.  
FT VARIANT 366 366 E -> D.  
FT VARIANT 398 398 Q -> K.  
FT VARIANT 419 419 K -> R.  
FT VARIANT 431 431 P -> R.  
FT VARIANT 439 439 T -> S.  
FT VARIANT 611 611 N -> I.  
FT VARIANT 706 706 I -> T.  
FT VARIANT 926 926 S -> T.  
FT VARIANT 952 952 E -> D.  
SQ SEQUENCE 977 AA; 109802 MW; FBF4AAE73BDF463B CRC64;  
  
Query Match 22.8%; Score 1198.5; DB 13; Length 977;  
Best Local Similarity 33.3%; Pred. No. 2e-96;  
Matches 325; Conservative 162; Mismatches 365; Indels 123; Gaps 33;  
  
QY 84 VAESGSITLQVQLATPGDLSCLWFKHSLGCGPFDLQNRGIV-----SMAILNVTE 136  
Db 37 IVDSGS---PLQLVCEGDAPVTEF-----PRLAKHKRYISKEVGKTRSFHVEKATV 84  
  
QY 137 TQAGEY-LLHIOSEIRANYTVLFTVNVRTQ-LYV-----LRRPYFRKMENQDALLCISEG 189  
Db 85 DFTGYKCVYINENDSNMSSVHVFRSDRVLFVSPSTSLR--YVRK-EGEDLLLPCLLT 141  
  
QY 190 VPEPTVENVLCSSHRESCKEKGPAVVRKEEVL-----HELFGTDIRCCARNALGRECTKL 245  
Db 142 DPDAT-DFTFRMDNGSRAAPYGMNATFDPKGLVIRNVHPGFNADYICSAICGAEKVSKT 200  
  
QY 246 FTIDLNOAQSTLPQLFLK-----VGEPLWIRCKAIHVNHGFLTWELEDKALEEGSY 298  
Db 201 FSNIIQRLRPP-PYVYLKRYNRVVKLVCEKQLQISCTNNPNFSYNTVTHSSRMLPKAE- 258  
  
QY 299 FEMSTYSTNRTMIRILLAFVSVSGRNDGTGYTCSSSKHPSQSALVT---ILEKGFINAT- 354



Db 259 -EKSTMEGDRLAIESILT-IPSVQQSDTGNICTGONEAGANSSTQQLLVVDEPYIRLSP 316  
QY 355 --SSQEEY-----EIDPYEKFCFSVRKAYPRIRC-----TWIESQASFCQEOR--GLEDG 401  
Db 317 KLSSKLTTHRGLSIEVSEGDVDLGVLEIAYPLTSHKWETPTSHNASLP--ENRFPYHNDR 375  
QY 402 YSISKFCDDHKN--KPGXYIFYAENDDAQFTKMTFLAIRKKPOVLAN--ASASQASCSDDGY 458  
Db 376 YEALLFLKRLNFEIGQYTLNVQNSMKNASITFDJMYTKPVAKVWENVVTTLSCPSYGY 435  
QY 459 PLPSWTWKCKSDKSNCTEIEI-----PEGVNNKANRKFVQWVSSSTLNMSEAGKGLLVK 514  
Db 436 PAPTILWYQCTGIRTCPCNTDLPQTQOTVEFKESFGAVGVESVLTVG--PNRRMTVV 494  
QY 515 CCAYNSMGWISCEIFLNSPQFPFFIODNISFYATIGLCLPPIV--VLIVLICHKKYKQFR 572  
Db 495 CVAFNLVGQSGDTFSMD-----VSDQITSAMCGSTVAMVVLGGLLIFMIYKXKQPR 547  
QY 573 YESOLOMIOVTGPDLDNEYFYVDFRDYEDLKWEPFRENLEFGKVLGSGAFGRVMNATAYG 632  
Db 548 YEIRWKIIIEATN--GNNYTFIDPTQLPYNERWEPPRDKLKLKTLGAGAFKVVVEATAYG 605  
QY 633 ISKTGVSIOAVAKMLKEKADSCKEALMSLKMTHLGHNDNIVNLLGACTLSGPVYLIF 692  
Db 606 LKREDNITRVAVKMLKASAPDEREALMSLKLILSHLGQHKNIIVNLLGACTHGGPVLVIT 665  
QY 693 EYCYGDLLNLYLSKRKFKHR-----TWEIFKEHNFSSYPTFAQHSNSSMPGS----- 741  
Db 666 EYCCHGDLNLFRLSKAENFLNFVMTIPNFPPEMTDYKNVSTIERMFVRSDSGISSTCSHY 725  
QY 742 ---REVQLHPPLDQLSGFNGNSIHSEDEIEYENQKRLAEEEEEDLNVLTFEDLLCFAYQV 798  
Db 726 LDMRPVTSRPTNSTLDS-----SSCOEDSWPLDMDDLLRFSSQV 765  
QY 799 AKGMEFLFKSCVHRDLAARNVLVTHGKVKIKCDFGLARDILSDSSYVVRGNARLPVKWM 858  
Db 766 AQGLDFLAAKNCIHROVAARNVLLTNSRVAKICDFGLAPDIMNDSNYVVGNGARLPVKWM 825  
QY 859 APESLFEIGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYKLIOSGFKMEQPPFYAT 918  
Db 826 APESIFECVTVQSDVWSYIMLWEIFSLGKSPYPNLIWDSKPYMKIKCGYOMSRPDPAP 885  
QY 919 EGIYFVMSQWAFDSRKRPSFPNLTSLGCOLAE--AEACIRTSTIHLPKAAPQOORGGLR 977  
Db 886 PEMTINKMCWNLDAAERPTFSKISQMIQRLMGETSEQDSQOEYKNIPAEAEQEQ--LE 943  
QY 978 AQSPQROVKIHRERS 992  
Db 944 SCDPAK----HEEES 954

Search completed: May 27, 2003, 14:34:27  
Job time : 83.4642 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: May 25, 2003, 02:59:08 ; Search time 8777.17 Seconds  
(without alignments)  
11608.404 Million cell updates/sec  
Title: US-09-919-408-3  
Perfect score: 3501  
Sequence: 1 CGAGCGGCATCCGAGGCT.....AAAAAAAAAAAAAAAAAAAA 3501

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :				GenEmbl :			
1:	gb_ba:	*		1:	gb_ba:	*	
2:	gb_htg:	*		2:	gb_htg:	*	
3:	gb_in:	*		3:	gb_in:	*	
4:	gb_om:	*		4:	gb_om:	*	
5:	gb_ov:	*		5:	gb_ov:	*	
6:	gb_pat:	*		6:	gb_pat:	*	
7:	gb_ph:	*		7:	gb_ph:	*	
8:	gb_pl:	*		8:	gb_pl:	*	
9:	gb_pr:	*		9:	gb_pr:	*	
10:	gb_ro:	*		10:	gb_ro:	*	
11:	gb_sts:	*		11:	gb_sts:	*	
12:	gb_sy:	*		12:	gb_sy:	*	
13:	gb_un:	*		13:	gb_un:	*	
14:	gb_vi:	*		14:	gb_vi:	*	
15:	em_ba:	*		15:	em_ba:	*	
16:	em_fun:	*		16:	em_fun:	*	
17:	em_hum:	*		17:	em_hum:	*	
18:	em_in:	*		18:	em_in:	*	
19:	em_mu:	*		19:	em_mu:	*	
20:	em_om:	*		20:	em_om:	*	
21:	em_or:	*		21:	em_or:	*	
22:	em_ov:	*		22:	em_ov:	*	
23:	em_pat:	*		23:	em_pat:	*	
24:	em_ph:	*		24:	em_ph:	*	
25:	em_pl:	*		25:	em_pl:	*	
26:	em_ro:	*		26:	em_ro:	*	
27:	em_sts:	*		27:	em_sts:	*	
28:	em_un:	*		28:	em_un:	*	
29:	em_vi:	*		29:	em_vi:	*	
30:	em_htg_hum:	*		30:	em_htg_hum:	*	
31:	em_htg_inv:	*		31:	em_htg_inv:	*	
32:	em_htg_other:	*		32:	em_htg_other:	*	
33:	em_htg_mus:	*		33:	em_htg_mus:	*	
34:	em_htg_pln:	*		34:	em_htg_pln:	*	
35:	em_htg_rnd:	*		35:	em_htg_rnd:	*	
36:	em_htg_mam:	*		36:	em_htg_mam:	*	
37:	em_htg_vrt:	*		37:	em_htg_vrt:	*	
38:	em_sy:	*		38:	em_sy:	*	
39:	em_htgo_hum:	*		39:	em_htgo_hum:	*	
40:	em_htgo_mus:	*		40:	em_htgo_mus:	*	
41:	em_htgo_other:	*		41:	em_htgo_other:	*	

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3501	100.0	3501	6	AR005212	AR005212 Sequence
2	3501	100.0	3501	6	AR071704	AR071704 Sequence
3	3501	100.0	3501	6	I25170	I25170 Sequence 3
4	3501	100.0	3501	6	I40601	I40601 Sequence 3
5	3475	99.3	3475	6	I44733	I44733 Sequence 3
6	3475	99.3	3475	6	HSU02687	U02687 Human growt
7	3474	99.2	3476	6	I40109	U02687 Human growt
8	3393.8	96.9	3422	9	HSFLT3RTK	Z26652 H.sapiens F
9	3083.2	88.1	3120	6	I44518	Z26652 H.sapiens F
10	2274.6	65.0	3521	6	I44732	I44518 Sequence 22
11	2274.6	65.0	3521	10	MMFLT3	I44732 Sequence 1
12	2227.2	63.6	3453	6	AR005211	X59398 Mouse Flt3
13	2227.2	63.6	3453	6	AR071703	AR005211 Sequence
14	2227.2	63.6	3453	6	I25169	AR071703 Sequence
15	2227.2	63.6	3453	6	I40600	I25169 Sequence 1
16	2203	62.9	3426	10	MUSFLK2	I40600 Sequence 1
17	624.4	17.8	1894	6	AR149571	M64689 Mouse flk-2
18	558.8	16.0	76721	9	AL591024	AR149571 Sequence
19	476.8	13.6	480	9	HUM3RTK	AL591024 Human DNA
20	400	11.4	400	11	G11126	L36162 Homo sapien
21	377.4	10.8	3545	5	AF153446	G11126 human STS S
22	336	9.6	3679	10	RRGSF1	AF153446 Danio rer
23	330.4	9.4	2894	4	ECA224643	X61479 Rat mRNA fo
24	330.4	9.4	2973	4	ECA224642	X61479 Rat mRNA fo
25	330.2	9.4	4577	14	FCSSMONC	AJ224643 Equus cab
26	329.8	9.4	3828	4	CATFMS	AJ224642 Equus cab
27	327.2	9.3	2894	4	ECA224644	K01643 Feline sarc
28	326.6	9.3	429	10	MUS3RTK	J03149 Cat (f.dome
29	325.6	9.3	2894	4	ECA224645	L36163 Mus musculu
30	324.8	9.3	2924	4	AF131209	AJ224644 Equus cab
31	324.6	9.3	2937	4	AF099030	AF131209 Trichosur
32	324.2	9.3	3992	6	AX394288	AF099030 Canis fam
33	324.2	9.3	3992	9	HSCFMS	AX394288 Sequence
34	323.2	9.2	2933	4	AF055037	X03663 Human mRNA
35	323.2	9.2	3575	10	MMCFCMS	AF055037 Equus cab
36	323	9.2	2979	4	AF448148	X68932 M.musculus
37	322.6	9.2	2176	4	AF263826	AF448148 Canis fam
38	322.6	9.2	2176	4	AF263827	AF263826 Bos tauru
39	322.6	9.2	3069	4	BOVCKR	AF263827 Bos tauru
40	322.2	9.2	3154	4	AF044249	D16680 Bovine mRNA
41	318.8	9.1	3826	5	XLXKRK1	AF044249 Canis fam
42	318.4	9.1	3665	10	MMFMSCR	Z48770 X.laavis c-
43	317.2	9.1	3771	4	GOTCKIT	X06368 Murine mRNA
44	317	9.1	5098	6	AX195906	D45168 Goat mRNA f
45	317	9.1	5098	10	MMCKIT	AX195906 Sequence
						Y00864 Mouse c-kit

ALIGNMENTS

RESULT 1	AR005212	AR005212	3501 bp	DNA	linear	PAT 04-DEC-1998
LOCUS	Sequence 3 from patent US 5747651.					
DEFINITION	Sequence 3 from patent US 5747651.					
ACCESSION	AR005212					
VERSION	AR005212.1	GI:3966091				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 3501)					
AUTHORS	Lemischka, I.R.					
TITLE	Antibodies against tyrosine kinase receptor flk-1					
JOURNAL	Patent: US 5747651-A 3 05-MAY-1998;					
FEATURES	Location/Qualifiers					



Db 2041 CTAAGATGATGACCCAGCTGGAGCCACGAGAAATATTGGAACCTGCTGGGGCGGTGC 2100  
QY 2101 ACATGTCAGGACCAATTTACTTGTGATTTTGAATACTGTGTGATGTGATCTTCTCAAC 2160  
Db 2101 ACATGTCAGGACCAATTTACTTGTGATTTTGAATACTGTGTGATGTGATCTTCTCAAC 2160  
QY 2161 TATCTAAGAAGTAAAGAGAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
Db 2161 TATCTAAGAAGTAAAGAGAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
QY 2221 AATTTCAGTTTATACCCACCTTCAATCACAATCCAAATTCAGCATGCCGTGTTCAAGA 2280  
Db 2221 AATTTCAGTTTATACCCACCTTCAATCACAATCCAAATTCAGCATGCCGTGTTCAAGA 2280  
QY 2281 GAAGTTCAGATACACCCGAGCTCGATCAAAATCTCAGGCTTCATGGGAATTCATTTCAC 2340  
Db 2281 GAAGTTCAGATACACCCGAGCTCGATCAAAATCTCAGGCTTCATGGGAATTCATTTCAC 2340  
QY 2341 TCTGAAGATGAATTTGAATATGAATCAAAACCAAAAGGCTGGAAGAGAGAGGACTTGAAT 2400  
Db 2341 TCTGAAGATGAATTTGAATATGAATCAAAACCAAAAGGCTGGAAGAGAGAGGACTTGAAT 2400  
QY 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTGCATATCAAGTTGCCAAAGGAATGAATTT 2460  
Db 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTGCATATCAAGTTGCCAAAGGAATGAATTT 2460  
QY 2461 CTGGAATTTAAGTCTGTGTTTACAGAGACCTGGCCGACGAAAGCTGCTTCACCCAC 2520  
Db 2461 CTGGAATTTAAGTCTGTGTTTACAGAGACCTGGCCGACGAAAGCTGCTTCACCCAC 2520  
QY 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGATGATCCAC 2580  
Db 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGATGATCCAC 2580  
QY 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTGATAATGATGGCCCGGAAAGCCGTGTTT 2640  
Db 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTGATAATGATGGCCCGGAAAGCCGTGTTT 2640  
QY 2641 GAAGGCATCTACACCATTAAGAGTATGCTGCTATGGAATATTTACTGTGGAAATTC 2700  
Db 2641 GAAGGCATCTACACCATTAAGAGTATGCTGCTATGGAATATTTACTGTGGAAATTC 2700  
QY 2701 TTCTCAGTTGGTGAATCTTACCTTGGCATTCGGTTGATGCTAACTTCTACAAACATG 2760  
Db 2701 TTCTCAGTTGGTGAATCTTACCTTGGCATTCGGTTGATGCTAACTTCTACAAACATG 2760  
QY 2761 ATTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATTATA 2820  
Db 2761 ATTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATTATA 2820  
QY 2821 ATGCAATCTGCTGGCTTTTACCTCAAGGAACGCCATCTTCCCTAATTTGACTTCG 2880  
Db 2821 ATGCAATCTGCTGGCTTTTACCTCAAGGAACGCCATCTTCCCTAATTTGACTTCG 2880  
QY 2881 TTTTATGAGTATGATGCTGGCAGATGCAAGAGGATGATATCAGAAATGGATGCGCT 2940  
Db 2881 TTTTATGAGTATGATGCTGGCAGATGCAAGAGGATGATATCAGAAATGGATGCGCT 2940  
QY 2941 GTTTCGGAATGCTCCTACACCTTACCAAAACAGCGACCTTTTACAGAGAGATGGAATTTG 3000  
Db 2941 GTTTCGGAATGCTCCTACACCTTACCAAAACAGCGACCTTTTACAGAGAGATGGAATTTG 3000  
QY 3001 GGGCTACTCTCCGAGGCTCAGGTGGAAGATTCGTAGAGAAACAATTTAGTTTAAAG 3060  
Db 3001 GGGCTACTCTCCGAGGCTCAGGTGGAAGATTCGTAGAGAAACAATTTAGTTTAAAG 3060  
QY 3061 ACTTCATCCCTCCACTATCCCTACAGGCTGTAGATTACCAAAACAAGATTAATTTCAAT 3120  
Db 3061 ACTTCATCCCTCCACTATCCCTACAGGCTGTAGATTACCAAAACAAGATTAATTTCAAT 3120  
QY 3121 CACTAAAGAAAATCTATTATCACTGCTTCCAGCAGCTTTTCTCTAGAGCGCTCT 3180

Db 3121 CACTAAAGAAAATCTATTATCACTGCTGCTTACCAGACTTTTCTCTAGAGCGCTCT 3180  
QY 3181 GCGTTTACTCTTGTGTTTCAAGGAGCTTTTGTAAATCAAAATCATCTGTCAAGAGCAG 3240  
Db 3181 GCGTTTACTCTTGTGTTTCAAGGAGCTTTTGTAAATCAAAATCATCTGTCAAGAGCAG 3240  
QY 3241 GAGGAGCTGATTAATGAACCTTTATTGGAGCATTTGATCTGATCAAGGCTTCTCAGGCCG 3300  
Db 3241 GAGGAGCTGATTAATGAACCTTTATTGGAGCATTTGATCTGATCAAGGCTTCTCAGGCCG 3300  
QY 3301 GCTTCAGTCAATTTGTGTAACCTGAAGTACAGTATATTCTTTGTAATACATAAAACAAAAGC 3360  
Db 3301 GCTTCAGTCAATTTGTGTAACCTGAAGTACAGTATATTCTTTGTAATACATAAAACAAAAGC 3360  
QY 3361 ATTTTGTCTAAGGAGAGCTTAATATGATTTTAAAGTCTATGTTTAAATTAATATGTAAG 3420  
Db 3361 ATTTTGTCTAAGGAGAGCTTAATATGATTTTAAAGTCTATGTTTAAATTAATATGTAAG 3420  
QY 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGGGAATAAAATTTTCTACTACAGAAAA 3480  
Db 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGGGAATAAAATTTTCTACTACAGAAAA 3480  
QY 3481 AAAAAAATAAAAAAATAAAAA 3501  
Db 3481 AAAAAAATAAAAAAATAAAAA 3501

RESULT 2  
AR071704  
LOCUS AR071704 3501 bp DNA linear PAT 18-FEB-2000  
DEFINITION Sequence 3 from patent US 5912133.  
ACCESSION AR071704  
VERSION AR071704.1 GI:7222592  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3501)  
AUTHORS Lemischka, I. R.  
TITLE Method for isolating stem cells expressing flk-1 receptors  
JOURNAL Patent: US 5912133-A 3 15-JUN-1999;  
FEATURES  
Location/Qualifiers  
Source l..3501  
BASE COUNT 1068 a 709 c 784 g 940 t  
ORIGIN

Query Match 100.0%; Score 3501; DB 6; Length 3501;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGAGGCGGATCCGAGGCTGGCGCGGCTGGGGGACCCCGGCTCCGGAGGCCATG 60  
Db 1 CGAGGCGGATCCGAGGCTGGCGCGGCTGGGGGACCCCGGCTCCGGAGGCCATG 60  
QY 61 CGCGGCTTGGCGCGACCGCGCACCGTGCCTGCTGCTGTTTCTGCAATGATA 120  
Db 61 CGCGGCTTGGCGCGACCGCGCACCGTGCCTGCTGCTGTTTCTGCAATGATA 120  
QY 121 TTTGGAGCTATTACAATAAAGATCTGCCCTGTGATCAAGTGTGTTTAAATCAATCAATAG 180  
Db 121 TTTGGAGCTATTACAATAAAGATCTGCCCTGTGATCAAGTGTGTTTAAATCAATCAATAG 180  
QY 181 ACAATGATTCATCAGTGGGAAGTATATCCCATGTTATCAGATCCCGGAA 240  
Db 181 ACAATGATTCATCAGTGGGAAGTATATCCCATGTTATCAGATCCCGGAA 240  
QY 241 GACCTCGGCTGTGCGTTGAGACCCAGAGCTCAGGAGCTGTACGAAGCTGCCGCTGTG 300  
Db 241 GACCTCGGCTGTGCGTTGAGACCCAGAGCTCAGGAGCTGTACGAAGCTGCCGCTGTG 300  
QY 301 GAAGTGGATGTATCTGCTTCATCAGACTGCAAGTGTGCTGATGCCCGGAGCAATT 360  
Db 301 GAAGTGGATGTATCTGCTTCATCAGACTGCAAGTGTGCTGATGCCCGGAGCAATT 360

Db 301 GAAAGTGATGTATCTGCTCCATCACACTGCAAGTGGTGGATGCCCCAGGAACATT 360  
Qy 361 TCCTGTCCTGGGTCCTTTAAGCACAGCTCCCTGAATTCGCCAGCCACATTTTGAATTACAA 420  
Db 361 TCCTGTCCTGGGTCCTTTAAGCACAGCTCCCTGAATTCGCCAGCCACATTTTGAATTACAA 420  
Qy 421 AACAGAGAGTGTGTTTCCATGTCATTTTGAATGACAGAAACCCAGCTGGAGAATAC 480  
Db 421 AACAGAGAGTGTGTTTCCATGTCATTTTGAATGACAGAAACCCAGCTGGAGAATAC 480  
Qy 481 CTACTTTTATTCAGAGTGAAGCTACCAATACACAATATTTTACAGTGAATAGA 540  
Db 481 CTACTTTTATTCAGAGTGAAGCTACCAATACACAATATTTTACAGTGAATAGA 540  
Qy 541 ATATCCCTGCTTACACATTAGAAGAGCTTACTTTAGAAAATGGAACCCAGAGCGCC 600  
Db 541 ATATCCCTGCTTACACATTAGAAGAGCTTACTTTAGAAAATGGAACCCAGAGCGCC 600  
Qy 601 CTGGTCTGCATATCTGAGAGCGTCCAGAGCGGATCGTGGATGGTGGTTCGATTCA 660  
Db 601 CTGGTCTGCATATCTGAGAGCGTCCAGAGCGGATCGTGGATGGTGGTTCGATTCA 660  
Qy 661 CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTTGTTAAAGAGAGGAAAAGTGCTT 720  
Db 661 CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGTTGTTAAAGAGAGGAAAAGTGCTT 720  
Qy 721 CATGAATATTGAGAGAGATAGGTGCTGCGAGAAATGAAGTGGAGGGAATGC 780  
Db 721 CATGAATATTGAGAGAGATAGGTGCTGCGAGAAATGAAGTGGAGGGAATGC 780  
Qy 781 ACCAGGCTGTTACCAATAGATCTAATCAAACTCTCGACACACATTTGCCAATATTT 840  
Db 781 ACCAGGCTGTTACCAATAGATCTAATCAAACTCTCGACACACATTTGCCAATATTT 840  
Qy 841 CTTAAAGTAGGGAAACCCCTTATGATAGGTGCAAGCTGTTTATGATGAACCATGATTC 900  
Db 841 CTTAAAGTAGGGAAACCCCTTATGATAGGTGCAAGCTGTTTATGATGAACCATGATTC 900  
Qy 901 GGGCTACCTGGGAATTAGAAACAAAGCACTCGAGGAGGCAACTCTTCAGATGAGT 960  
Db 901 GGGCTACCTGGGAATTAGAAACAAAGCACTCGAGGAGGCAACTCTTCAGATGAGT 960  
Qy 961 ACCTATTCAACAAACAGAACTATGATACGATTTCTGTTGCTTTGTTATCATCAGTGGCA 1020  
Db 961 ACCTATTCAACAAACAGAACTATGATACGATTTCTGTTGCTTTGTTATCATCAGTGGCA 1020  
Qy 1021 AGAAACGACACCGGATACACTGTTCTCTTCAAAGCATCCCAAGTCAATCAGCTTTG 1080  
Db 1021 AGAAACGACACCGGATACACTGTTCTCTTCAAAGCATCCCAAGTCAATCAGCTTTG 1080  
Qy 1081 GTTACCATCGTAGGAAGGGATTTAATATGCTACCAATTCAGTGAAGATTATGAATTT 1140  
Db 1081 GTTACCATCGTAGGAAGGGATTTAATATGCTACCAATTCAGTGAAGATTATGAATTT 1140  
Qy 1141 GACCAATATGAAGATTTGTTTCTGTCAGTCTTAAAGCTACCCACAAATCAGATGT 1200  
Db 1141 GACCAATATGAAGATTTGTTTCTGTCAGTCTTAAAGCTACCCACAAATCAGATGT 1200  
Qy 1201 ACGTGGACCTTCTCTCGAAAATCATTTCTGTTGAGCAAAAGGGTCTTGATAACGGATAC 1260  
Db 1201 ACGTGGACCTTCTCTCGAAAATCATTTCTGTTGAGCAAAAGGGTCTTGATAACGGATAC 1260  
Qy 1261 AGCATATCCAAGTTTTCGAATCATAGACAGCCAGGAGATATATATTCATGCGAGAA 1320  
Db 1261 AGCATATCCAAGTTTTCGAATCATAGACAGCCAGGAGATATATATTCATGCGAGAA 1320  
Qy 1321 AATGATGATGCCAATTTACCAAAATGTTACGCTCAATATAGAAGAGAAACCTCAAGTG 1380  
Db 1321 AATGATGATGCCAATTTACCAAAATGTTACGCTCAATATAGAAGAGAAACCTCAAGTG 1380  
Qy 1381 CTCGAGAAAGCATCGGCAAGTCAGGCGTCTGTTTCTCGGATGATACCCATTTACCATCT 1440  
Db 1381 CTCGAGAAAGCATCGGCAAGTCAGGCGTCTGTTTCTCGGATGATACCCATTTACCATCT 1440

Qy 1441 TGGACCTGGAAGAAGTGTTCAGACAAGTCTCCAACTGCACAGAAGAGATCACAGAAGGA 1500  
Db 1441 TGGACCTGGAAGAAGTGTTCAGACAAGTCTCCAACTGCACAGAAGAGATCACAGAAGGA 1500  
Qy 1501 GTCTGGAATAGAAAGGCTAACAGAAAAGTGTGTTGGACAGTGGGTGTCGAGCAGTACTCTA 1560  
Db 1501 GTCTGGAATAGAAAGGCTAACAGAAAAGTGTGTTGGACAGTGGGTGTCGAGCAGTACTCTA 1560  
Qy 1561 ACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACATAATTCCTTGGC 1620  
Db 1561 ACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACATAATTCCTTGGC 1620  
Qy 1621 ACATCTGTGAGACCATCCTTTTAAACTCTCCAGGCCCCCTTCCCTTTTATCCAAAGCAAC 1680  
Db 1621 ACATCTGTGAGACCATCCTTTTAAACTCTCCAGGCCCCCTTCCCTTTTATCCAAAGCAAC 1680  
Qy 1681 ATCTCATTTATGCAACAATTTGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
Db 1681 ATCTCATTTATGCAACAATTTGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
Qy 1741 ATTTGTCACAAGTACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Db 1741 ATTTGTCACAAGTACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Qy 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTGGGAAGTACTAGATCAGTGCTTT 1860  
Db 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTGGGAAGTACTAGATCAGTGCTTT 1860  
Qy 1861 AAATGGGAGTTTCCAAAGAAAATTTAGAGTTTGGGAAGTACTAGATCAGTGCTTTT 1920  
Db 1861 AAATGGGAGTTTCCAAAGAAAATTTAGAGTTTGGGAAGTACTAGATCAGTGCTTTT 1920  
Qy 1921 GAAAAGTATGAACCCACAGCTTATGGAATTTAGCAAAACAGAGTCTCAATCCAGGTT 1980  
Db 1921 GAAAAGTATGAACCCACAGCTTATGGAATTTAGCAAAACAGAGTCTCAATCCAGGTT 1980  
Qy 1981 GCGCTCAAAATGCTGAAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTTCAGAA 2040  
Db 1981 GCGCTCAAAATGCTGAAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTTCAGAA 2040  
Qy 2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGCGTGC 2100  
Db 2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGCGTGC 2100  
Qy 2101 ACACCTCAGGACCAATTTTACTTGTATTTTGAATGATGTTGCTATGTTGATCTCTCAAC 2160  
Db 2101 ACACCTCAGGACCAATTTTACTTGTATTTTGAATGATGTTGCTATGTTGATCTCTCAAC 2160  
Qy 2161 TATCTAAGAGTAAAAGAGAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
Db 2161 TATCTAAGAGTAAAAGAGAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
Qy 2221 AATTTTCAGTGTTCACCCACTTTTCCAAATCATCTCAAAATTCAGAGTGCCTGGTTCAGA 2280  
Db 2221 AATTTTCAGTGTTCACCCACTTTTCCAAATCATCTCAAAATTCAGAGTGCCTGGTTCAGA 2280  
Qy 2281 GAAGTTCAGATACACCCGACTCGGATCAATCTCAGGCTTTCATGGGAATTCATTTTCAAC 2340  
Db 2281 GAAGTTCAGATACACCCGACTCGGATCAATCTCAGGCTTTCATGGGAATTCATTTTCAAC 2340  
Qy 2341 TCTGAAGATGAATTCGAATATGAAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
Db 2341 TCTGAAGATGAATTCGAATATGAAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
Qy 2401 GTGCTTACATTTTGAAGATCTCTTTGCTTTGCTATCAAGTTCGCAAGGAATGGAATTT 2460  
Db 2401 GTGCTTACATTTTGAAGATCTCTTTGCTTTGCTATCAAGTTCGCAAGGAATGGAATTT 2460  
Qy 2461 CTGGAATTTAAGTCTGTTTTCAGAGACCTGGCCCGCCAGGACGCTGCTTGTACCCAC 2520  
Db 2461 CTGGAATTTAAGTCTGTTTTCAGAGACCTGGCCCGCCAGGACGCTGCTTGTACCCAC 2520

Qy	2521	GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTTCCAAC	2580
Db	2521	GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTTCCAAC	2580
Qy	2581	TATGTTGTCAGGGCAATGCCGCTCGTCGCCGTAAAATGGATGGCCCCGAGAAAGCCTGTGTTT	2640
Db	2581	TATGTTGTCAGGGCAATGCCGCTCGTCGCCGTAAAATGGATGGCCCCGAGAAAGCCTGTGTTT	2640
Qy	2641	GAAGGCATCTACACCAATTAAGAGTGATGTCGTGTCATATGGAATATTTACTGTGGGAAATC	2700
Db	2641	GAAGGCATCTACACCAATTAAGAGTGATGTCGTGTCATATGGAATATTTACTGTGGGAAATC	2700
Qy	2701	TTCTCACTTGGTGTGAATCCTTACCCTGGCATTCGCCGTTGATGCTACTACTCTACAAACTG	2760
Db	2701	TTCTCACTTGGTGTGAATCCTTACCCTGGCATTCGCCGTTGATGCTACTACTCTACAAACTG	2760
Qy	2761	ATTCAAATGGATTTAAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATTTATA	2820
Db	2761	ATTCAAATGGATTTAAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATTTATA	2820
Qy	2821	ATGCAATCTCTGGGGCTTTTGACTCAAGGAAACGGGCATCTTCGCTAAATTTGACTTCG	2880
Db	2821	ATGCAATCTCTGGGGCTTTTGACTCAAGGAAACGGGCATCTTCGCTAAATTTGACTTCG	2880
Qy	2881	TTTTTTAGTAGTGCAGCTGGCAGATGCAGAGAGAGCGATGTATCAGATGTGGATGGCCGT	2940
Db	2881	TTTTTTAGTAGTGCAGCTGGCAGATGCAGAGAGAGCGATGTATCAGATGTGGATGGCCGT	2940
Qy	2941	GTTTTCGGAATGTCCTCACACCTACCAAACAGGCGACCTTTTCAGCAGAGATGGATTTG	3000
Db	2941	GTTTTCGGAATGTCCTCACACCTACCAAACAGGCGACCTTTTCAGCAGAGATGGATTTG	3000
Qy	3001	GGGCTACTCTCTCGCAGGCTCAGTCTGGAAGATTGCTAGAGGAACAATTTAGTTTAAAGG	3060
Db	3001	GGGCTACTCTCTCGCAGGCTCAGTCTGGAAGATTGCTAGAGGAACAATTTAGTTTAAAGG	3060
Qy	3061	ACTTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAACCAAGATTAAATTCAT	3120
Db	3061	ACTTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAACCAAGATTAAATTCAT	3120
Qy	3121	CACATAAGAAAATCTATTTATCAACTGCTGCTTCCACAGACTTTTCTCTAGAAGCCGTCT	3180
Db	3121	CACATAAGAAAATCTATTTATCAACTGCTGCTTCCACAGACTTTTCTCTAGAAGCCGTCT	3180
Qy	3181	GCCTTTACTCTCTTTCAAAGGACTTTTGTAAATCAATCATCCTGTCACAAGGCAG	3240
Db	3181	GCCTTTACTCTCTTTCAAAGGACTTTTGTAAATCAATCATCCTGTCACAAGGCAG	3240
Qy	3241	GAGGAGCTGATAATGAACCTTTATTGGAGCATTGATCTGCATCCAAGGCCCTCTCAGGCCG	3300
Db	3241	GAGGAGCTGATAATGAACCTTTATTGGAGCATTGATCTGCATCCAAGGCCCTCTCAGGCCG	3300
Qy	3301	GCCTTGAGTGAATGTGTACCTGAAGTACAGTATATCTTCTGTAATACATAAAACAAAGC	3360
Db	3301	GCCTTGAGTGAATGTGTACCTGAAGTACAGTATATCTTCTGTAATACATAAAACAAAGC	3360
Qy	3361	ATTTTTCGTAAGGAGAGCTAAATGATTTTTTTTAAAGTCTATGTTTTAAATTAATATGTAAA	3420
Db	3361	ATTTTTCGTAAGGAGAGCTAAATGATTTTTTTTAAAGTCTATGTTTTAAATTAATATGTAAA	3420
Qy	3421	TTTTTTCAGCTATTATTAGTGATATATTTTATGGGTGGGAATAAAATTTCTACTACAGAAAAA	3480
Db	3421	TTTTTTCAGCTATTATTAGTGATATATTTTATGGGTGGGAATAAAATTTCTACTACAGAAAAA	3480
Qy	3481	AAAAAAAAAAAAAAAAAAAAA 3501	
Db	3481	AAAAAAAAAAAAAAAAAAAAA 3501	

RESULT 3			
I25170			
LOCUS	I25170	3501 bp	DNA
DEFINITION	Sequence 3 from patent US 5548065.		
		linear	PAT 07-OCT-1996

ACCESSION	I25170
VERSION	I25170.1 GI:1605040
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 3501)
TITLE	Lemishka,I.R.
JOURNAL	Tyrosine kinase receptor human flk-2-specific antibodies
FEATURES	Patent: us 5548065-A 3 20-AUG-1996;
source	Location/Qualifiers
BASE COUNT	1..3501
ORIGIN	/organism="unknown"
	1068 a 709 c 784 g 940 t
Query Match	100.0%; Score 3501; DB 6; Length 3501;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 3501:	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 CGAGCCGGCATCCGAGGGCTGGGCCGGCGCCCTGGGGGACC CCGGGGTCCGGAGGCCATG 60 
Dd	1 CGAGCCGGCATCCGAGGGCTGGGCCGGCGCCCTGGGGGACC CCGGGGTCCGGAGGCCATG 60 
Oy	61 CCGCGTTGGCGCGCAGCGGGCACCGTGCCTGCTGTGTTTTCTGCAATGATA 120 
Dd	61 CCGCGTTGGCGCGCAGCGGGCACCGTGCCTGCTGTGTTTTCTGCAATGATA 120 
Oy	121 TTTGGGACTATTACAATCAAGATCTGCCCTGTGATCAAGTGTTTTTAATCAATCATAA 180 
Dd	121 TTTGGGACTATTACAATCAAGATCTGCCCTGTGATCAAGTGTTTTTAATCAATCATAA 180 
Oy	181 AACAAATGATTCATCAGTGGGAAGTCAATCATATCCCATGTATCAGAATCCCCGGAA 240 
Dd	181 AACAAATGATTCATCAGTGGGAAGTCAATCATATCCCATGTATCAGAATCCCCGGAA 240 
Oy	241 GACCTCGGGTGTGGCTTTAGACACCCAGAGCTCAGGAGACAGTGTACGAAGTGCCTGTG 300 
Dd	241 GACCTCGGGTGTGGCTTTAGACACCCAGAGCTCAGGAGACAGTGTACGAAGTGCCTGTG 300 
Oy	301 GAAGTGATGTAICTGCTTCCATCACACTGCAAGTGCTGCTGATGCCCGGAGAACATT 360 
Dd	301 GAAGTGATGTAICTGCTTCCATCACACTGCAAGTGCTGCTGATGCCCGGAGAACATT 360 
Oy	361 TCCTGCTCTCGGGTCTTTAAGCAGAGCTCCCTGAATTGCCAGGCACATTTGATTACAA 420 
Dd	361 TCCTGCTCTCGGGTCTTTAAGCAGAGCTCCCTGAATTGCCAGGCACATTTGATTACAA 420 
Oy	421 AACAGAGGAGTGTGTTCCATGGTCAATTTGAAAATGACAGAAACCCCAAGCTGGAGAA 480 
Dd	421 AACAGAGGAGTGTGTTCCATGGTCAATTTGAAAATGACAGAAACCCCAAGCTGGAGAA 480 
Oy	481 CTACTTTTTATTACAGATGAAGTACCATTACACAATATTTTACAGTGAAGTATAAGA 540 
Dd	481 CTACTTTTTATTACAGATGAAGTACCATTACACAATATTTTACAGTGAAGTATAAGA 540 
Oy	541 AATACCCCTGCTTTACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACCCAGGAGCC 600 
Dd	541 AATACCCCTGCTTTACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACCCAGGAGCC 600 
Oy	601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTTCGATTCA 660 
Dd	601 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTTCGATTCA 660 
Oy	661 CAGGGGAAAGCTGTAAAGAAAGATPCCAGCTGTGTTTAAAGGAGGAAAAAGTGCCT 720 
Dd	661 CAGGGGAAAGCTGTAAAGAAAGATPCCAGCTGTGTTTAAAGGAGGAAAAAGTGCCT 720 
Oy	721 CATGAATTTATTGGGACGGACATAAGTGCTGTGCCAGAAATGAACTGGCGAGGAATGC 780 
Dd	721 CATGAATTTATTGGGACGGACATAAGTGCTGTGCCAGAAATGAACTGGCGAGGAATGC 780 
Oy	781 ACCAGGCTGTTCAAAATAGATCTAAATCAAACCTCCTCAGACCCACATTTGCCACAATATTT 840 

Db 781 ACCAGGCTGTTACAAATAGATCTAAATCAAACCTCTCAGACACATATGCCACAATATATT 840  
Qy 841 CTTAAAGTAGGGAAACCCCTTATGGATTAAGGTCAAAGCTGTTCATGTGAACCATGGATTC 900  
Db 841 CTTAAAGTAGGGAAACCCCTTATGGATTAAGGTCAAAGCTGTTCATGTGAACCATGGATTC 900  
Qy 901 GGGCTCACCCTGGGAATTAGAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT 960  
Db 901 GGGCTCACCCTGGGAATTAGAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT 960  
Qy 961 ACCTATTCAACAAACAGAACTATGATACGGATTCCTTTGCTTTGTATCATCAGTGGCA 1020  
Db 961 ACCTATTCAACAAACAGAACTATGATACGGATTCCTTTGCTTTGTATCATCAGTGGCA 1020  
Qy 1021 AGAAACGACACCCGGATCTACACTTGTCTCTCAAAGCATCCCAAGTCAATCAGCTTG 1080  
Db 1021 AGAAACGACACCCGGATCTACACTTGTCTCTCAAAGCATCCCAAGTCAATCAGCTTG 1080  
Qy 1081 GTTACCATCGTAGGAAAGGGATTTATAAATGCTACCAATTCAAGTGAAGATTATGAAT 1140  
Db 1081 GTTACCATCGTAGGAAAGGGATTTATAAATGCTACCAATTCAAGTGAAGATTATGAAT 1140  
Qy 1141 GACCAATATGAAGATTTGTTTTCTGTCTCAGGTTTAAAGCCTACCCACAATATCAGATG 1200  
Db 1141 GACCAATATGAAGATTTGTTTTCTGTCTCAGGTTTAAAGCCTACCCACAATATCAGATG 1200  
Qy 1201 ACGTGGACCTCTCTCGAAAATCAATTCCTCTGTGACAAAAGGGTCTTGATACAGGATAC 1260  
Db 1201 ACGTGGACCTCTCTCGAAAATCAATTCCTCTGTGACAAAAGGGTCTTGATACAGGATAC 1260  
Qy 1261 AGCATATCCAAGTTTGGCAATCATTAAGCACAGCCAGGAGATATATATTCCATGCAAG 1320  
Db 1261 AGCATATCCAAGTTTGGCAATCATTAAGCACAGCCAGGAGATATATATTCCATGCAAG 1320  
Qy 1321 AATGATGATGCCAATTTACCAAAATGTTACGCTGAATATTAAGAGGAAACCTCAAGTG 1380  
Db 1321 AATGATGATGCCAATTTACCAAAATGTTACGCTGAATATTAAGAGGAAACCTCAAGTG 1380  
Qy 1381 CTCGCAAGCATTCGGCAAGTCAGGCGTCTCTTCTCGGATGGATACCCATTACCATCT 1440  
Db 1381 CTCGCAAGCATTCGGCAAGTCAGGCGTCTCTTCTCGGATGGATACCCATTACCATCT 1440  
Qy 1441 TGGACCTGGAAGAAGTGTTCAGACAAGTCTCCAACTGCACAGAAGATCACAAGGA 1500  
Db 1441 TGGACCTGGAAGAAGTGTTCAGACAAGTCTCCAACTGCACAGAAGATCACAAGGA 1500  
Qy 1501 GTCTGGAATAGAAAGCTTAACAGAAAAGTGTGGACAGTGGGTGTCGAGCAGTACTCTA 1560  
Db 1501 GTCTGGAATAGAAAGCTTAACAGAAAAGTGTGGACAGTGGGTGTCGAGCAGTACTCTA 1560  
Qy 1561 AACATGAGTGAAGCCATAAAAGGGTTCCTGGTCAAGTGTGTCATACAAATTCCTTGGC 1620  
Db 1561 AACATGAGTGAAGCCATAAAAGGGTTCCTGGTCAAGTGTGTCATACAAATTCCTTGGC 1620  
Qy 1621 ACATCTGTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAACAGACAAC 1680  
Db 1621 ACATCTGTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAACAGACAAC 1680  
Qy 1681 ATCTCATTTATGCAACAATTTGGTGTGTTGTCCTCTCTTCTCATTTGTCGTTTAAACCCCTGCTA 1740  
Db 1681 ATCTCATTTATGCAACAATTTGGTGTGTTGTCCTCTCTTCTCATTTGTCGTTTAAACCCCTGCTA 1740  
Qy 1741 ATTTGTACAAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800  
Db 1741 ATTTGTACAAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800  
Qy 1801 ACCGCTCCTCAGATAATCAGTACTTCTACGTTGATTTCCAGAGAATATCAATATGATCTC 1860  
Db 1801 ACCGCTCCTCAGATAATCAGTACTTCTACGTTGATTTCCAGAGAATATCAATATGATCTC 1860  
Qy 1861 AAATGGGAGTTTCCAAAGGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
Db 1861 AAATGGGAGTTTCCAAAGGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920

Db 1861 AAATGGGAGTTTCCAAAGGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
Qy 1921 GGAAAAGTGAAGCAACAGCTTATGGAATTTAGCAAAACAGAGAGTCTCAATCCAGGTT 1980  
Db 1921 GGAAAAGTGAAGCAACAGCTTATGGAATTTAGCAAAACAGAGAGTCTCAATCCAGGTT 1980  
Qy 1981 GCCTGCAAAATGCTGTAAGAAAAGCAGACAGCTCTGAAAAGAGAGGAGCTCATGTCTCAGAA 2040  
Db 1981 GCCTGCAAAATGCTGTAAGAAAAGCAGACAGCTCTGAAAAGAGAGGAGCTCATGTCTCAGAA 2040  
Qy 2041 CTCAGATGATCACCAGCTGGAGCCAGCAAGATATTTGTGAACCTGCTGGGGGGCTGC 2100  
Db 2041 CTCAGATGATCACCAGCTGGAGCCAGCAAGATATTTGTGAACCTGCTGGGGGGCTGC 2100  
Qy 2101 ACACGTCTCAGGACCAATTTACTTTGATTTTGAATCTGTTGCTATGCTGATCTCTCAAC 2160  
Db 2101 ACACGTCTCAGGACCAATTTACTTTGATTTTGAATCTGTTGCTATGCTGATCTCTCAAC 2160  
Qy 2161 TATCTAAGAAGTAAAGAGAAAATTTTCAGAGACTTTGACAGAGATTTTCAAGGAACAC 2220  
Db 2161 TATCTAAGAAGTAAAGAGAAAATTTTCAGAGACTTTGACAGAGATTTTCAAGGAACAC 2220  
Qy 2221 AATTTTCAGTTTACCCCACTTTCCCAATCACATCCAAATTCACAGCATGCTGCTGTTCAAGA 2280  
Db 2221 AATTTTCAGTTTACCCCACTTTCCCAATCACATCCAAATTCACAGCATGCTGCTGTTCAAGA 2280  
Qy 2281 GAAGTTCAGATACACCGGACTCGGATCAAACTCAGGGCTTCATGGGAATTCATTTCAC 2340  
Db 2281 GAAGTTCAGATACACCGGACTCGGATCAAACTCAGGGCTTCATGGGAATTCATTTCAC 2340  
Qy 2341 TCTGAAGATGAATTTGAATATGAAAACCAAAAAGGCTGGAAGAAGAGGAGGACTTGAAT 2400  
Db 2341 TCTGAAGATGAATTTGAATATGAAAACCAAAAAGGCTGGAAGAAGAGGAGGACTTGAAT 2400  
Qy 2401 GTGCTTACATTTCAAGATCTCTTTCTGCTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
Db 2401 GTGCTTACATTTCAAGATCTCTTTCTGCTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
Qy 2461 CTGGAATTTAAGTCGTGTGTTTACAGAGACTGCGCGCCAGGAAACGCTGTGTCACCCAC 2520  
Db 2461 CTGGAATTTAAGTCGTGTGTTTACAGAGACTGCGCGCCAGGAAACGCTGTGTCACCCAC 2520  
Qy 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTCGAGATATCATGAGTGAATTCACAC 2580  
Db 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTCGAGATATCATGAGTGAATTCACAC 2580  
Qy 2581 TATGTTGTGAGGGCAATGCCCTCTGCCCTGTAATGATGGATGCGCCCGGAAAGCCCTCTT 2640  
Db 2581 TATGTTGTGAGGGCAATGCCCTCTGCCCTGTAATGATGGATGCGCCCGGAAAGCCCTCTT 2640  
Qy 2641 GAAGGCATCTACACCATTAAAGATGATGCTGCTCATATGCAATATTTACTGTGGGAATC 2700  
Db 2641 GAAGGCATCTACACCATTAAAGATGATGCTGCTCATATGCAATATTTACTGTGGGAATC 2700  
Qy 2701 TTCTCACTTTGGTGTGAATCCTTTACCCTGGCATTTCGGTTGATGCTAACCTTCTACAACTG 2760  
Db 2701 TTCTCACTTTGGTGTGAATCCTTTACCCTGGCATTTCGGTTGATGCTAACCTTCTACAACTG 2760  
Qy 2761 ATTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTTACAGAAAGAAATATACATTATA 2820  
Db 2761 ATTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTTACAGAAAGAAATATACATTATA 2820  
Qy 2821 ATGCAATCTGCTGGGCTTTTGACTCAAGGAAAGCCCATCTTCCCTAATTTGACTTCG 2880  
Db 2821 ATGCAATCTGCTGGGCTTTTGACTCAAGGAAAGCCCATCTTCCCTAATTTGACTTCG 2880  
Qy 2881 TTTTAAAGTGTGAGTGGCAGATGCAAGAAAGCAGATGATCAGATGTTGGATGGCGGT 2940  
Db 2881 TTTTAAAGTGTGAGTGGCAGATGCAAGAAAGCAGATGATCAGATGTTGGATGGCGGT 2940  
Qy 2941 GTTTCGAATGCTCTCACACCTTACCACCAAGAGCGGACCTTTTACAGAGAGATGGATTG 3000  
Db 2941 GTTTCGAATGCTCTCACACCTTACCACCAAGAGCGGACCTTTTACAGAGAGATGGATTG 3000

```
QY 3001 GGGCTACTCTCTCCGAGGCTCAGCTCGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG 3060
|||||
DB 3001 GGGCTACTCTCTCCGAGGCTCAGCTCGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG 3060
|||||
QY 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAAGATTAATTTTCAT 3120
|||||
DB 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAAGATTAATTTTCAT 3120
|||||
QY 3121 CACTAAAAGAAAATCTATTATCACTGCTGCTTACAGAGACTTTTCTCTAGAACCGGCTCT 3180
|||||
DB 3121 CACTAAAAGAAAATCTATTATCACTGCTGCTTACAGAGACTTTTCTCTAGAACCGGCTCT 3180
|||||
QY 3181 CGCTTTTACTCTGTTTTCAAAAGGAGCTTTTGTAAAATCAAAATCATCTGTCACAAAGGAG 3240
|||||
DB 3181 CGCTTTTACTCTGTTTTCAAAAGGAGCTTTTGTAAAATCAAAATCATCTGTCACAAAGGAG 3240
|||||
QY 3241 GAGGAGCTGATAATGAACCTTTATTGGAGCATTGATCTGCATCCAGAGGCTTCTCAGGCGG 3300
|||||
DB 3241 GAGGAGCTGATAATGAACCTTTATTGGAGCATTGATCTGCATCCAGAGGCTTCTCAGGCGG 3300
|||||
QY 3301 GCTTGAGTGAATTTGTGTACCTGAAGTACAGTATATTTTGAAGTCTATGTTTAAATATATGTAAA 3420
|||||
DB 3301 GCTTGAGTGAATTTGTGTACCTGAAGTACAGTATATTTTGAAGTCTATGTTTAAATATATGTAAA 3420
|||||
QY 3421 TTTTTCAGCTATTAGTATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAAA 3480
|||||
DB 3421 TTTTTCAGCTATTAGTATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAAA 3480
|||||
QY 3481 AAAAAAATAAAAAAAAAAAAAA 3501
|||||
DB 3481 AAAAAAATAAAAAAAAAAAAAA 3501
|||||

RESULT 4
140601
LOCUS 140601 3501 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 3 from patent US 5621090.
ACCESSION I40601
VERSION I40601.1 GI:2082893
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3501)
AUTHORS Lemischka,I.R.
TITLE Nucleic acids encoding soluble human FLK-2 extracellular domain
JOURNAL Patent: US 5621090-A 3 15-APR-1997;
FEATURES Location/Qualifiers
source 1..3501
/organism="unknown"
BASE COUNT 1068 a 709 c 784 g 940 t
ORIGIN
Query Match 100.0%; Score 3501; DB 6; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGAGCGGCATCCGAGGCTGGCGCGGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60
|||||
DB 1 CGAGCGGCATCCGAGGCTGGCGCGGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60
|||||
QY 61 CCGGGTGGCGCGGACCGGGACCGCTGGCGTCTGTTTCTTCTGCAATGATA 120
|||||
DB 61 CCGGGTGGCGCGGACCGGGACCGCTGGCGTCTGTTTCTTCTGCAATGATA 120
|||||
QY 121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAAG 180
|||||
DB 121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAAG 180
|||||
```

```
QY 181 AACAAATGATTTCATCAGTGGGAAGTTCATCATATATCCCATGTCATCAGAAATCCCGGAA 240
|||||
DB 181 AACAAATGATTTCATCAGTGGGAAGTTCATCATATATCCCATGTCATCAGAAATCCCGGAA 240
|||||
QY 241 GACCTCGGCTGTGCGTTGAGACCCAGAGCTCAGGACAGAGTGTACGAAGTGCCTGCTGTG 300
|||||
DB 241 GACCTCGGCTGTGCGTTGAGACCCAGAGCTCAGGACAGAGTGTACGAAGTGCCTGCTGTG 300
|||||
QY 301 GAAGTGGATGTATCTGCTTCCATCACACTGCAAGTCTGGTTCATGCCCCAGGAACATTT 360
|||||
DB 301 GAAGTGGATGTATCTGCTTCCATCACACTGCAAGTCTGGTTCATGCCCCAGGAACATTT 360
|||||
QY 361 TCCTGCTCTCTGCTCTTTTAAAGCACAGCTCCCTCAATTTGCCAGGCACATTTTGAATTTACAA 420
|||||
DB 361 TCCTGCTCTCTGCTCTTTTAAAGCACAGCTCCCTCAATTTGCCAGGCACATTTTGAATTTACAA 420
|||||
QY 421 AACAGAGGAGTTGTTTCCATGTTTGAATAATGACAGAAACCCCAAGCTGGAGATAC 480
|||||
DB 421 AACAGAGGAGTTGTTTCCATGTTTGAATAATGACAGAAACCCCAAGCTGGAGATAC 480
|||||
QY 481 CTACTTTTATTTACAGAGTGAAGCTACCAATTTACACAATATTTGTTTACAGTGAATAAGA 540
|||||
DB 481 CTACTTTTATTTACAGAGTGAAGCTACCAATTTACACAATATTTGTTTACAGTGAATAAGA 540
|||||
QY 541 AATACCTGCTTTTACACATTTAAGAACCTTACTTTAGAAAATGGAACACAGGAGGCC 600
|||||
DB 541 AATACCTGCTTTTACACATTTAAGAACCTTACTTTAGAAAATGGAACACAGGAGGCC 600
|||||
QY 601 CTGGTCTGCATATCTGAGAGCGTTCAGAGCGCATGCTGGAATGGTGTTCCTCCGATTCA 660
|||||
DB 601 CTGGTCTGCATATCTGAGAGCGTTCAGAGCGCATGCTGGAATGGTGTTCCTCCGATTCA 660
|||||
QY 661 CAGGGGAAAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTTAAAAAGGAGGAAAAAGTCTT 720
|||||
DB 661 CAGGGGAAAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTTAAAAAGGAGGAAAAAGTCTT 720
|||||
QY 721 CATGAATATTGTCGACGACATAGGTGCTGTCAGCAAAATGAATGCGCAGGGAATGC 780
|||||
DB 721 CATGAATATTGTCGACGACATAGGTGCTGTCAGCAAAATGAATGCGCAGGGAATGC 780
|||||
QY 781 ACCAGGCTGTTCCACAATAGATCTAAATCAAACTCCTCAGACCACTTGCACAAATATTT 840
|||||
DB 781 ACCAGGCTGTTCCACAATAGATCTAAATCAAACTCCTCAGACCACTTGCACAAATATTT 840
|||||
QY 841 CTTAAAGTAGGGGAACCCCTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCATGGATTC 900
|||||
DB 841 CTTAAAGTAGGGGAACCCCTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCATGGATTC 900
|||||
QY 901 GGGCTCACCTGGGAATTAGAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT 960
|||||
DB 901 GGGCTCACCTGGGAATTAGAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT 960
|||||
QY 961 ACCTATTCAACAAACAGAACTATGATACGATTCTGTTTGTCTTTTATCATCAGTGGCA 1020
|||||
DB 961 ACCTATTCAACAAACAGAACTATGATACGATTCTGTTTGTCTTTTATCATCAGTGGCA 1020
|||||
QY 1021 AGAAACGACACCGGATACACTTGTTCCTCTCAAAGCATCCAGTCAATCAGCTTTG 1080
|||||
DB 1021 AGAAACGACACCGGATACACTTGTTCCTCTCAAAGCATCCAGTCAATCAGCTTTG 1080
|||||
QY 1081 GTTACCATCGTAGGAAGGATTTTATAATGCTACCAATTCAGTGAAGATTAATGAAAT 1140
|||||
DB 1081 GTTACCATCGTAGGAAGGATTTTATAATGCTACCAATTCAGTGAAGATTAATGAAAT 1140
|||||
QY 1141 GACCAATATGAGAGTTTGTCTGTGAGTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT 1200
|||||
DB 1141 GACCAATATGAGAGTTTGTCTGTGAGTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT 1200
|||||
QY 1201 AGCTGGAGCTTCTCTCGAAAATCATTTCCCTTGTGAGCAAAAGGCTTGTGATAACGATAC 1260
|||||
DB 1201 AGCTGGAGCTTCTCTCGAAAATCATTTCCCTTGTGAGCAAAAGGCTTGTGATAACGATAC 1260
|||||
```



1261 AGCATATCCAAAGTTTGCATATCAATAAGCACAGCCAGGAGAAATATATATTCATGCGAGAA 1320  
1261 AGCATATCCAAAGTTTGCATATCAATAAGCACAGCCAGGAGAAATATATATTCATGCGAGAA 1320  
1321 AATGATGATGCCAAATTTACCAAAATGTTACGCTGAATATAAGAAGAAACCTCAAGTG 1380  
1321 AATGATGATGCCAAATTTACCAAAATGTTACGCTGAATATAAGAAGAAACCTCAAGTG 1380  
1381 CTCGCAAGGATCGGCAAGTCAGCGCTCCTGTTTCTCGGATGGATACCCATTACCATTCT 1440  
1381 CTCGCAAGGATCGGCAAGTCAGCGCTCCTGTTTCTCGGATGGATACCCATTACCATTCT 1440  
1441 TGGACCTTGGAAAGAGTGTTCAGAAAGTCTCCCAACTGCACAGAAGAGATCACAGAAGGA 1500  
1501 GTCTGGAAATAGAAAGGCTTAACAGAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
1501 GTCTGGAAATAGAAAGGCTTAACAGAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
1561 AACATGAGTGAAGCCATAAAGAGGTTCTGCTCAAGTGTCTGCATACAATTCCTCTGGC 1620  
1561 AACATGAGTGAAGCCATAAAGAGGTTCTGCTCAAGTGTCTGCATACAATTCCTCTGGC 1620  
1621 ACATCTTGTGAGAGATCCTTTTAACTCTCCAGGCCCTTCCCTTTTCAATCAAGACAAC 1680  
1621 ACATCTTGTGAGAGATCCTTTTAACTCTCCAGGCCCTTCCCTTTTCAATCAAGACAAC 1680  
1681 ATCTCATTTATGCAACAATTTGGTGTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
1681 ATCTCATTTATGCAACAATTTGGTGTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
1741 ATTTGTCAACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
1741 ATTTGTCAACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
1801 ACCGCTCCTCAGATTAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860  
1801 ACCGCTCCTCAGATTAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860  
1861 AAATGGGAGTTTCCAAAGAGAAATTTAGATTTGGGAAGGTACTAGGATCAGGTCTTTT 1920  
1861 AAATGGGAGTTTCCAAAGAGAAATTTAGATTTGGGAAGGTACTAGGATCAGGTCTTTT 1920  
1921 GGAAAGTGTACGACGCAACGCTTATGGAATAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
1921 GGAAAGTGTACGACGCAACGCTTATGGAATAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
1981 GCGGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
1981 GCGGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
2041 CTCAGATGATGACCCAGCTGGAGGCCAGGAGATATTTGTAACCTGCTGGGGGCGTGC 2100  
2041 CTCAGATGATGACCCAGCTGGAGGCCAGGAGATATTTGTAACCTGCTGGGGGCGTGC 2100  
2101 ACAGTGTACAGGACCAATTTACTTTGATTTTGAATAGTGTGATGATGATGATGATGATGAT 2160  
2101 ACAGTGTACAGGACCAATTTACTTTGATTTTGAATAGTGTGATGATGATGATGATGATGAT 2160  
2161 TATCTAAGAGTAAAAGAGAAAATTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
2161 TATCTAAGAGTAAAAGAGAAAATTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
2221 AATTTCAAGTAAAAGAGAAAATTTTCAATCCAAATTTCCAGCATGCTGGTGTCAAGA 2280  
2221 AATTTCAAGTAAAAGAGAAAATTTTCAATCCAAATTTCCAGCATGCTGGTGTCAAGA 2280  
2281 GAAGTTTCAGATACACCGGACTCGGATCAATCTCAGGCTTCAATGGGAATTCATTTTCA 2340  
2281 GAAGTTTCAGATACACCGGACTCGGATCAATCTCAGGCTTCAATGGGAATTCATTTTCA 2340  
2341 TCTGAAGATGAAATTCGAATATGAAACCAAAAAAGGCTGGAAGAAGAGGAGGACTTGAAT 2400

2341 TCTGAAGATGAAATTCGAATATGAAACCAAAAAAGGCTGGAAGAAGAGGAGGACTTGAAT 2400  
2401 GTGCTTACATTTGAAGATCTTTTGTGTTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
2401 GTGCTTACATTTGAAGATCTTTTGTGTTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
2461 CTGGAATTTAAGTCTGTGTTTACAGAGACCTGGCCGCGCAGAACGCTGCTGTGTCACCCAC 2520  
2461 CTGGAATTTAAGTCTGTGTTTACAGAGACCTGGCCGCGCAGAACGCTGCTGTGTCACCCAC 2520  
2521 GGGAAAGTGTGAAGATATGTGACTTTTGGATTGCTCGAGATATATCATGAGTGAATCCAAC 2580  
2521 GGGAAAGTGTGAAGATATGTGACTTTTGGATTGCTCGAGATATATCATGAGTGAATCCAAC 2580  
2581 TATGTTGTACAGGGCAATGCCCGTCTGCTGTAAAATGGATGGCCCCGGAAGACCTGTTT 2640  
2581 TATGTTGTACAGGGCAATGCCCGTCTGCTGTAAAATGGATGGCCCCGGAAGACCTGTTT 2640  
2641 GAAGGCATCTACACCATTAAGAGTGTGCTGCTCATATGGAATATATCTGTGGGAATC 2700  
2641 GAAGGCATCTACACCATTAAGAGTGTGCTGCTCATATGGAATATATCTGTGGGAATC 2700  
2701 TTCTCACTTGGTGTGAATCCTTACCTTGGCATTCGGGTTGATGCTAACTTCTACAACTG 2760  
2701 TTCTCACTTGGTGTGAATCCTTACCTTGGCATTCGGGTTGATGCTAACTTCTACAACTG 2760  
2761 ATTTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTTACAGAGAATAATACATTATA 2820  
2761 ATTTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTTACAGAGAATAATACATTATA 2820  
2821 ATGCAATCTGCTGGGCTTTTACCTCAAGAAACGCCCATCTTCCCTCAATTTGACTTCG 2880  
2821 ATGCAATCTGCTGGGCTTTTACCTCAAGAAACGCCCATCTTCCCTCAATTTGACTTCG 2880  
2881 TTTTGTAGATGTCAGTGGCAGATGCAAGAAGCAGATGATCAGAAATGTTGATGGGCTG 2940  
2881 TTTTGTAGATGTCAGTGGCAGATGCAAGAAGCAGATGATCAGAAATGTTGATGGGCTG 2940  
2941 GTTTCGGAATGTCCTCACACCTTACCAAAACAGCGGACCTTTTACAGAGAGATGGATTTG 3000  
2941 GTTTCGGAATGTCCTCACACCTTACCAAAACAGCGGACCTTTTACAGAGAGATGGATTTG 3000  
3001 GGGCTACTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060  
3001 GGGCTACTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060  
3061 ACTTCATCTCCACCTATCCCTTAACAGGCTGTAGATTACCAAAACAGATTAATTTTCAT 3120  
3061 ACTTCATCTCCACCTATCCCTTAACAGGCTGTAGATTACCAAAACAGATTAATTTTCAT 3120  
3121 CACTAAAAGAAAATCTATTATCAACTGCTGCTTACAGAGACTTTTCTCTAGAAGCGCTCT 3180  
3121 CACTAAAAGAAAATCTATTATCAACTGCTGCTTACAGAGACTTTTCTCTAGAAGCGCTCT 3180  
3181 GGGTTTACTCTGTTTTCAAAAGGACTTTTGTAAAATCAATCATCTGTCACAGGCGAG 3240  
3181 GGGTTTACTCTGTTTTCAAAAGGACTTTTGTAAAATCAATCATCTGTCACAGGCGAG 3240  
3241 GAGGAGCTGATATGAACCTTTATGGAGCATTCATCTGCAATCAAGGCTTCTCAGGCGG 3300  
3241 GAGGAGCTGATATGAACCTTTATGGAGCATTCATCTGCAATCAAGGCTTCTCAGGCGG 3300  
3301 GCTTGAAGTAAATGTGTACCTGAAGTACAGTATATCTTCTTAATACATAAAACAAAGC 3360  
3301 GCTTGAAGTAAATGTGTACCTGAAGTACAGTATATCTTCTTAATACATAAAACAAAGC 3360  
3361 ATTTTGTGAAGGAGAGCTAATATGATTTTTTAAAGTCTATGTTTTTAAATATATGTA 3420  
3361 ATTTTGTGAAGGAGAGCTAATATGATTTTTTAAAGTCTATGTTTTTAAATATATGTA 3420  
3421 TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAATTTCTACTACAGAAAA 3480

Db 3421 TTTTTCAGCTATTAGTCATATATTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480

Qy 3481 AAAAAAAAAAAAAAAAAAAAA 3501

Db 3481 AAAAAAAAAAAAAAAAAAAAA 3501

RESULT 5

LOCUS I44733 3475 bp DNA linear PAT 07-OCT-1997

DEFINITION Sequence 3 from patent US 5635388.

ACCESSION I44733

VERSION I44733.1 GI:2469446

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 3475)

AUTHORS Bennett,B.D., Broz,S.D., Matthews,W. and Zeigler,F.C.

TITLE Agonist antibodies against the flk2/flt3 receptor and uses thereof

JOURNAL Patent: US 5635388-A 3 03-JUN-1997;

FEATURES

Location/Qualifiers

1..3475

/organism="unknown"

BASE COUNT 1042 a 709 c 784 g 940 t

ORIGIN

Query Match 99.3%; Score 3475; DB 6; Length 3475;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGAGCGGCATCCAGAGCTGGCGCCGCGCTGGGGACCCCGGGCTCCGGAGGCCATG 60

Db 1 CGAGCGGCATCCAGAGCTGGCGCCGCGCTGGGGACCCCGGGCTCCGGAGGCCATG 60

Qy 61 CGGCGTTGGCGCGCAGCGGGCACCGTGGCTGCTGTTTTCGCAATGATA 120

Db 61 CGGCGTTGGCGCGCAGCGGGCACCGTGGCTGCTGTTTTCGCAATGATA 120

Qy 121 TTGGGACTATTACAATCAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAAG 180

Db 121 TTGGGACTATTACAATCAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAAG 180

Qy 181 AACAAATGATCATCAGTGGGGAAGTCATCATATCCATGGTATCAGATCCCGGAA 240

Db 181 AACAAATGATCATCAGTGGGGAAGTCATCATATCCATGGTATCAGATCCCGGAA 240

Qy 241 GACCTCGGGTGTGGTTGAGACCCAGAGCTCAGGACAGTGTACGAACTGCCGCTGTG 300

Db 241 GACCTCGGGTGTGGTTGAGACCCAGAGCTCAGGACAGTGTACGAACTGCCGCTGTG 300

Qy 301 GAAGTGGATGATCTGCTTCCATCAGCTCAAGTGTGCTGATGCCCCAGGGAACATT 360

Db 301 GAAGTGGATGATCTGCTTCCATCAGCTCAAGTGTGCTGATGCCCCAGGGAACATT 360

Qy 361 TCCTGTCTCGGCTTTAAGCAGAGCTCCCTGAATGCCAGCCACATTTGATTACAA 420

Db 361 TCCTGTCTCGGCTTTAAGCAGAGCTCCCTGAATGCCAGCCACATTTGATTACAA 420

Qy 421 AACAGAGAGTGTGTTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC 480

Db 421 AACAGAGAGTGTGTTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC 480

Qy 481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTTACAGTGAATATA 540

Db 481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTTACAGTGAATATA 540

Qy 541 AATACCTGCTTTACACATTAAGAGAGCTTACTTTAGAAAAATGGAAAAACCCAGCACC 600

Db 541 AATACCTGCTTTACACATTAAGAGAGCTTACTTTAGAAAAATGGAAAAACCCAGCACC 600

Qy 601 CTGCTGTCATATCTGAGAGGTTCCAGAGCGGATCGTGGAAATGGTGGCTTTTCCGATTCA 660

Db 601 CTGCTGTCATATCTGAGAGGTTCCAGAGCGGATCGTGGAAATGGTGGCTTTTCCGATTCA 660

Qy 661 CAGGGGAAAGCTGTAAAGAAAGAACTCCAGCTGTGTTTAAAGAGGAGGAAAGTGCCT 720

Db 661 CAGGGGAAAGCTGTAAAGAAAGAACTCCAGCTGTGTTTAAAGAGGAGGAAAGTGCCT 720

Qy 721 CATGAATTTATTTGGGACGACATAAGTGTCTGCAGAAATGAACCTGGCGAGGAATGC 780

Db 721 CATGAATTTATTTGGGACGACATAAGTGTCTGCAGAAATGAACCTGGCGAGGAATGC 780

Qy 781 ACCAGGCTGTTTCAAAATAGATCTAAATCAAACTCTCAGACACATTTGACAAATTTT 840

Db 781 ACCAGGCTGTTTCAAAATAGATCTAAATCAAACTCTCAGACACATTTGACAAATTTT 840

Qy 841 CTTAAAGTAGGGGAACCCCTTATGGATAGGTGCAAGCTGTTTCATGTGAACCATGATTC 900

Db 841 CTTAAAGTAGGGGAACCCCTTATGGATAGGTGCAAGCTGTTTCATGTGAACCATGATTC 900

Qy 901 GGGCTCACTGGGAATTAGAAACAAAGCACTCGAGGGGCAACTTCTTGGATGAGT 960

Db 901 GGGCTCACTGGGAATTAGAAACAAAGCACTCGAGGGGCAACTTCTTGGATGAGT 960

Qy 961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTTTTATCATCATCAGTGC 1020

Db 961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTTTTATCATCATCAGTGC 1020

Qy 1021 AGAAACGACCGGATACACTTGTTCCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080

Db 1021 AGAAACGACCGGATACACTTGTTCCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080

Qy 1081 GTTACCATCTGAGGAAGGATTTATATATGCTACCAATTCAGTGAAGATATGAAAT 1140

Db 1081 GTTACCATCTGAGGAAGGATTTATATATGCTACCAATTCAGTGAAGATATGAAAT 1140

Qy 1141 GACCAATATGAGAGTGTGTTTCTGTGAGTTTAAAGCCTACCCACAAATCAGATGT 1200

Db 1141 GACCAATATGAGAGTGTGTTTCTGTGAGTTTAAAGCCTACCCACAAATCAGATGT 1200

Qy 1201 ACGTGGACCTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTCTATACGGATAC 1260

Db 1201 ACGTGGACCTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTCTATACGGATAC 1260

Qy 1261 AGCATATCAAGTGTGTTTGAATCATTAAGCACCGAGGAGATATATATTCATCGAGAA 1320

Db 1261 AGCATATCAAGTGTGTTTGAATCATTAAGCACCGAGGAGATATATATTCATCGAGAA 1320

Qy 1321 AATGATGATGCCAATTTACCAAAATGTTTCAGCTGAATATAGAAGAGAACTCAAGTG 1380

Db 1321 AATGATGATGCCAATTTACCAAAATGTTTCAGCTGAATATAGAAGAGAACTCAAGTG 1380

Qy 1381 CTCGAGAGCATCGGCAAGTCCAGGCTCTGTTCTCGATGGATACCATTCATCT 1440

Db 1381 CTCGAGAGCATCGGCAAGTCCAGGCTCTGTTCTCGATGGATACCATTCATCT 1440

Qy 1441 TGGACCTGGAAGAGTGTTCAGCAAGTCTCCCACTGCACAGAGAGATCAGAGAGGA 1500

Db 1441 TGGACCTGGAAGAGTGTTCAGCAAGTCTCCCACTGCACAGAGAGATCAGAGAGGA 1500

Qy 1501 GTCTGGAATAGAAAGGCTTAACAGAAAGTGTGACAGTGGGTGTCGAGCAGTACTCTA 1560

Db 1501 GTCTGGAATAGAAAGGCTTAACAGAAAGTGTGACAGTGGGTGTCGAGCAGTACTCTA 1560

Qy 1561 AACATGATGAGCCATAAAGGTTCTGCTCAAGTGTGTCATACAAATCCCTTGGC 1620

Db 1561 AACATGATGAGCCATAAAGGTTCTGCTCAAGTGTGTCATACAAATCCCTTGGC 1620

Qy 1621 ACATCTTGTGAGACCATCTTTTAACTCTCCAGGCCCCCTTCCCTTTCACAGACAAC 1680

Db 1621 ACATCTTGTGAGACCATCTTTTAACTCTCCAGGCCCCCTTCCCTTTCACAGACAAC 1680

Qy 1681 ATCTCATTTCTATGCAACAAATTTGTTGCTCTCTTTCATTTGCTTTTAAACCTGCTA 1740

Db 1681 ATCTCATTTCTATGCAACAAATTTGTTGCTCTCTTTCATTTGCTTTTAAACCTGCTA 1740

601 CTGCTGTCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGTGGCTTTTCCGATTCA 660

661 CAGGGGAAAGCTGTAAAGAAAGAACTCCAGCTGTGTTTAAAGAGGAGGAAAGTGCCT 720

661 CAGGGGAAAGCTGTAAAGAAAGAACTCCAGCTGTGTTTAAAGAGGAGGAAAGTGCCT 720

721 CATGAATTTATTTGGGACGACATAAGTGTCTGCAGAAATGAACCTGGCGAGGAATGC 780

721 CATGAATTTATTTGGGACGACATAAGTGTCTGCAGAAATGAACCTGGCGAGGAATGC 780

781 ACCAGGCTGTTTCAAAATAGATCTAAATCAAACTCTCAGACACATTTGACAAATTTT 840

781 ACCAGGCTGTTTCAAAATAGATCTAAATCAAACTCTCAGACACATTTGACAAATTTT 840

841 CTTAAAGTAGGGGAACCCCTTATGGATAGGTGCAAGCTGTTTCATGTGAACCATGATTC 900

841 CTTAAAGTAGGGGAACCCCTTATGGATAGGTGCAAGCTGTTTCATGTGAACCATGATTC 900

901 GGGCTCACTGGGAATTAGAAACAAAGCACTCGAGGGGCAACTTCTTGGATGAGT 960

901 GGGCTCACTGGGAATTAGAAACAAAGCACTCGAGGGGCAACTTCTTGGATGAGT 960

961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTTTTATCATCATCAGTGC 1020

961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTGCTTTTATCATCATCAGTGC 1020

1021 AGAAACGACCGGATACACTTGTTCCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080

1021 AGAAACGACCGGATACACTTGTTCCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080

1081 GTTACCATCTGAGGAAGGATTTATATATGCTACCAATTCAGTGAAGATATGAAAT 1140

1081 GTTACCATCTGAGGAAGGATTTATATATGCTACCAATTCAGTGAAGATATGAAAT 1140

1141 GACCAATATGAGAGTGTGTTTCTGTGAGTTTAAAGCCTACCCACAAATCAGATGT 1200

1141 GACCAATATGAGAGTGTGTTTCTGTGAGTTTAAAGCCTACCCACAAATCAGATGT 1200

1201 ACGTGGACCTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTCTATACGGATAC 1260

1201 ACGTGGACCTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTCTATACGGATAC 1260

1261 AGCATATCAAGTGTGTTTGAATCATTAAGCACCGAGGAGATATATATTCATCGAGAA 1320

1261 AGCATATCAAGTGTGTTTGAATCATTAAGCACCGAGGAGATATATATTCATCGAGAA 1320

1321 AATGATGATGCCAATTTACCAAAATGTTTCAGCTGAATATAGAAGAGAACTCAAGTG 1380

1321 AATGATGATGCCAATTTACCAAAATGTTTCAGCTGAATATAGAAGAGAACTCAAGTG 1380

1381 CTCGAGAGCATCGGCAAGTCCAGGCTCTGTTCTCGATGGATACCATTCATCT 1440

1381 CTCGAGAGCATCGGCAAGTCCAGGCTCTGTTCTCGATGGATACCATTCATCT 1440

1441 TGGACCTGGAAGAGTGTTCAGCAAGTCTCCCACTGCACAGAGAGATCAGAGAGGA 1500

1441 TGGACCTGGAAGAGTGTTCAGCAAGTCTCCCACTGCACAGAGAGATCAGAGAGGA 1500

1501 GTCTGGAATAGAAAGGCTTAACAGAAAGTGTGACAGTGGGTGTCGAGCAGTACTCTA 1560

1501 GTCTGGAATAGAAAGGCTTAACAGAAAGTGTGACAGTGGGTGTCGAGCAGTACTCTA 1560

1561 AACATGATGAGCCATAAAGGTTCTGCTCAAGTGTGTCATACAAATCCCTTGGC 1620

1561 AACATGATGAGCCATAAAGGTTCTGCTCAAGTGTGTCATACAAATCCCTTGGC 1620

1621 ACATCTTGTGAGACCATCTTTTAACTCTCCAGGCCCCCTTCCCTTTCACAGACAAC 1680

1621 ACATCTTGTGAGACCATCTTTTAACTCTCCAGGCCCCCTTCCCTTTCACAGACAAC 1680

1681 ATCTCATTTCTATGCAACAAATTTGTTGCTCTCTTTCATTTGCTTTTAAACCTGCTA 1740

1681 ATCTCATTTCTATGCAACAAATTTGTTGCTCTCTTTCATTTGCTTTTAAACCTGCTA 1740

1741 ATTTGTCACAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGTTACAGGTG 1800  
1741 ATTTGTCACAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGTTACAGGTG 1800  
1801 ACCGGCTCCTCAGATAATGAGTACTCTACGTTGATTTTCAGAGATATGAATATCATCTC 1860  
1801 ACCGGCTCCTCAGATAATGAGTACTCTACGTTGATTTTCAGAGATATGAATATCATCTC 1860  
1861 AAATGGGAGTTTCCAAAGAGAAATTTAGATTTGGGAAGGTACTAGGATCAGGTCTTTT 1920  
1861 AAATGGGAGTTTCCAAAGAGAAATTTAGATTTGGGAAGGTACTAGGATCAGGTCTTTT 1920  
1921 GGAAAGTGTGAACGCAACGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
1921 GGAAAGTGTGAACGCAACGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
1981 GCGGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTTCAGAA 2040  
1981 GCGGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTTCAGAA 2040  
2041 CTCAAGATGATCACCAGCTGGAGGACGAGAGATATTTGTGAACCTGCTGGGGCGTGC 2100  
2041 CTCAAGATGATCACCAGCTGGAGGACGAGAGATATTTGTGAACCTGCTGGGGCGTGC 2100  
2101 ACACGTCTCAGGACCAATTTACTGTTGATTTTGAATCTGTTGATGTTGATCTTCTCAAC 2160  
2101 ACACGTCTCAGGACCAATTTACTGTTGATTTTGAATCTGTTGATGTTGATCTTCTCAAC 2160  
2161 TATCTAAGAAGTAAAGAGAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
2161 TATCTAAGAAGTAAAGAGAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
2221 AATTTCAAGTTTACCAGCTTTTCAATCAGATCAATTCAGATGCTGCTGTTCAAGA 2280  
2221 AATTTCAAGTTTACCAGCTTTTCAATCAGATCAATTCAGATGCTGCTGTTCAAGA 2280  
2281 GAAGTTCAGATACACCCGACCTCGATCAATCTCAGGCTTCATGGGAATTCATTTCAC 2340  
2281 GAAGTTCAGATACACCCGACCTCGATCAATCTCAGGCTTCATGGGAATTCATTTCAC 2340  
2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAGAGGAGGACCTTGAAT 2400  
2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAGAGGAGGACCTTGAAT 2400  
2401 GTGCTTACATTTGAAGATCTTCTGCTTGGATATCAAGTTGCCAAAGGAATGGAATTT 2460  
2401 GTGCTTACATTTGAAGATCTTCTGCTTGGATATCAAGTTGCCAAAGGAATGGAATTT 2460  
2461 CTGGAATTTAAGTCTGTTTCCAGAGACCTGGCGCCAGGAACGCTGTTGTCACCCAC 2520  
2461 CTGGAATTTAAGTCTGTTTCCAGAGACCTGGCGCCAGGAACGCTGTTGTCACCCAC 2520  
2521 GGGAAAGTGGTGAAGATATGTGACATTTGGATTTGGCTCGAGATATCATGATTTCCAC 2580  
2521 GGGAAAGTGGTGAAGATATGTGACATTTGGATTTGGCTCGAGATATCATGATTTCCAC 2580  
2581 TATGTTGTCAGGGCAATGCCGCTGCTGTAATGATGGCCCCCGGAAAGCCCTGTTT 2640  
2581 TATGTTGTCAGGGCAATGCCGCTGCTGTAATGATGGCCCCCGGAAAGCCCTGTTT 2640  
2641 GAAGCATCTACACCATTAAGAGTATGCTGCTCATATGGAATATTTACTGTGGGAATC 2700  
2641 GAAGCATCTACACCATTAAGAGTATGCTGCTCATATGGAATATTTACTGTGGGAATC 2700  
2701 TTCTCACTTGGTGTGAATCCTTACCCTGGCATTCGGTTGATGCTCAACTTCAAACTG 2760  
2701 TTCTCACTTGGTGTGAATCCTTACCCTGGCATTCGGTTGATGCTCAACTTCAAACTG 2760  
2761 ATTCAAAATGGATTTAAATGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820  
2761 ATTCAAAATGGATTTAAATGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820

2821 ATGCAATCTGCTGGGCTTTTGAATCAAGAAACAGCCATCTCTCCCTAATTTGACTTCG 2880  
2821 ATGCAATCTGCTGGGCTTTTGAATCAAGAAACAGCCATCTCTCCCTAATTTGACTTCG 2880  
2881 TTTTGTAGTATGCTGCTGCGATGTCAGAGAGAGCATGATATCAGATGTTGATGCGCGT 2940  
2881 TTTTGTAGTATGCTGCTGCGATGTCAGAGAGAGCATGATATCAGATGTTGATGCGCGT 2940  
2941 GTTTCGGATGCTCTCACACCTTACCAAAACAGCGGACCTTTTTCAGCAGAGAGATGGAATTG 3000  
2941 GTTTCGGATGCTCTCACACCTTACCAAAACAGCGGACCTTTTTCAGCAGAGAGATGGAATTG 3000  
3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGAACCAATTTAGTTTAAAG 3060  
3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGAACCAATTTAGTTTAAAG 3060  
3061 ACTTCATCTCTCCCTACCTATCCCTACAGGCTGTAGATTTACCAAAACAGGATTAATTTTCAT 3120  
3061 ACTTCATCTCTCCCTACCTATCCCTACAGGCTGTAGATTTACCAAAACAGGATTAATTTTCAT 3120  
3121 CACTAAAGAAAATCTATTATCAACTGCTGCTTCAGCAGACTTTTCTTGAAGCCGCTCT 3180  
3121 CACTAAAGAAAATCTATTATCAACTGCTGCTTCAGCAGACTTTTCTTGAAGCCGCTCT 3180  
3181 GCGTTTACTCTGTTTTCAAAGGAGCTTTTGTAAAATCAATCATCTGTCAAGAGCAG 3240  
3181 GCGTTTACTCTGTTTTCAAAGGAGCTTTTGTAAAATCAATCATCTGTCAAGAGCAG 3240  
3241 GAGGAGCTATATGAATGAATTTTATGAGGATTCATCTGCTCAAGGCTTCTCAGGCGG 3300  
3241 GAGGAGCTATATGAATGAATTTTATGAGGATTCATCTGCTCAAGGCTTCTCAGGCGG 3300  
3301 GCTTGAGTGAATTTGCTACCTGAGTACAGTATTTCTTAAATACATATAAACAAGC 3360  
3301 GCTTGAGTGAATTTGCTACCTGAGTACAGTATTTCTTAAATACATATAAACAAGC 3360  
3361 ATTTGCTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAAAATATATGTA 3420  
3361 ATTTGCTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAAAATATATGTA 3420  
3421 TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAAATTTCTACTACAG 3475  
3421 TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAAATTTCTACTACAG 3475

## RESULT 6

HSU02687  
LOCUS  
DEFINITION  
Human growth factor receptor tyrosine kinase (STK-1) mRNA, complete cds.

ACCESSION  
VERSION  
KEYWORDS  
U02687  
U02687.1 GI:409572

SOURCE  
ORGANISM  
human.  
Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
Small, D., Levenstein, M., Kim, E., Carow, C., Amin, S., Rockwell, P., Witte, L., Burrow, C., Ratajczak, M. Z., Gewirtz, A. M. et al. STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in CD34+ human bone marrow cells and is involved in the proliferation of early progenitor/stem cells  
Proc. Natl. Acad. Sci. U.S.A. 91 (2), 459-463 (1994)

JOURNAL  
MEDLINE  
PUBMED  
94119906  
7507245

REFERENCE  
2 (bases 1 to 3475)

AUTHORS  
Small, D.

TITLE  
Direct Submission

Submitted (25-OCT-1993) Donald Small, Oncology, Johns Hopkins University School of Medicine, 600 N. Wolfe St., Baltimore, MD 21287 USA

Location/Qualifiers

## FEATURES





Matches 3475; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	CGAGCGGCATCCAGAGGCTGGCGCGCCCTGGGGAGACCCGGGCTCCGGAGGCCATG	60
Db	1	CGAGCGGCATCCAGAGGCTGGCGCGCCCTGGGGAGACCCGGGCTCCGGAGGCCATG	60
Qy	61	CCGCGGTGGCGCGAGCGGGGACCGGTGCCGTGCTGCTGTTTTTCTGCAATGATA	120
Db	61	CCGCGGTGGCGCGAGCGGGGACCGGTGCCGTGCTGCTGTTTTTCTGCAATGATA	120
Qy	121	TTTGGGACTATTACAATCAAGATCTGCCGTGATCAAGTGTGTTTTTAATCAATCAAG	180
Db	121	TTTGGGACTATTACAATCAAGATCTGCCGTGATCAAGTGTGTTTTTAATCAATCAAG	180
Qy	181	AACAATGATTCATCAGTGGGAAGTCATCATATCCATGGTATCAGATCCCGGAA	240
Db	181	AACAATGATTCATCAGTGGGAAGTCATCATATCCATGGTATCAGATCCCGGAA	240
Qy	241	GACCTCGGGTGGTGGTGGAGCCCGAGAGCTCAGGACAGTGTACGAAGCTGCCGTGTG	300
Db	241	GACCTCGGGTGGTGGTGGAGCCCGAGAGCTCAGGACAGTGTACGAAGCTGCCGTGTG	300
Qy	301	GAAGTGTATCTGCTTCCATCACACTGCAAGTGTGCTGATGCCCGAGGAACATT	360
Db	301	GAAGTGTATCTGCTTCCATCACACTGCAAGTGTGCTGATGCCCGAGGAACATT	360
Qy	361	TCCTGTCTGGGCTTTAAGCACAGCTCCCTGAATGCCAGCCACATTTTGATTACAA	420
Db	361	TCCTGTCTGGGCTTTAAGCACAGCTCCCTGAATGCCAGCCACATTTTGATTACAA	420
Qy	421	AACAGAGAGTGTGTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC	480
Db	421	AACAGAGAGTGTGTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC	480
Qy	481	CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTGTACAGTCAATAGA	540
Db	481	CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTGTACAGTCAATAGA	540
Qy	541	AATACCCCTGTTTACATTAAGAGACCTTACTTTAGAAAATGGAACACGAGCGCC	600
Db	541	AATACCCCTGTTTACATTAAGAGACCTTACTTTAGAAAATGGAACACGAGCGCC	600
Qy	601	CTGCTCTGCATATCTGAGAGCGTTCAGAGCCGATCGTGGATGGTCTTGGCATCA	660
Db	601	CTGCTCTGCATATCTGAGAGCGTTCAGAGCCGATCGTGGATGGTCTTGGCATCA	660
Qy	661	CAGGGGAAAGCTTAAGAGAAAGTCCAGCTGTTGTTAAAGAGGAGGAAAAGTGCTT	720
Db	661	CAGGGGAAAGCTTAAGAGAAAGTCCAGCTGTTGTTAAAGAGGAGGAAAAGTGCTT	720
Qy	721	CATGAATTTTGGGACGGACATAGGTGCTGTGCCAGAAATGAACCTGGGAGGAATGC	780
Db	721	CATGAATTTTGGGACGGACATAGGTGCTGTGCCAGAAATGAACCTGGGAGGAATGC	780
Qy	781	ACCAGGCTGTTTCAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCACAATATT	840
Db	781	ACCAGGCTGTTTCAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCACAATATT	840
Qy	841	CTTAAAGTAGGGAAACCTTATGGATTAAGTGTGCAAGCTGTTTCAATCAATGAGTTC	900
Db	841	CTTAAAGTAGGGAAACCTTATGGATTAAGTGTGCAAGCTGTTTCAATCAATGAGTTC	900
Qy	901	GGGCTACCTGGGAAATAGAAAACAAAGCACTCAGGAGGCACTACTTTGAGATGAGT	960
Db	901	GGGCTACCTGGGAAATAGAAAACAAAGCACTCAGGAGGCACTACTTTGAGATGAGT	960
Qy	961	ACCTATTCAACAAACAACTATCATACGGATTCTGTTGCTTTGTATCATCATAGTSCA	1020
Db	961	ACCTATTCAACAAACAACTATCATACGGATTCTGTTGCTTTGTATCATCATAGTSCA	1020
Qy	1021	AGAAACGACCGGATACACTTGTTCCTTCAAGAGATCCAGTCAATCAGCTTGG	1080
Db	1021	AGAAACGACCGGATACACTTGTTCCTTCAAGAGATCCAGTCAATCAGCTTGG	1080

Qy	1081	GTTACCATCGTAGGAAAGGGATTTATAATGCTACCAATTCAGTGAAGATTATGAAATT	1140
Db	1081	GTTACCATCGTAGGAAAGGGATTTATAATGCTACCAATTCAGTGAAGATTATGAAATT	1140
Qy	1141	GACCAATATGAAGAGATTTTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCATGT	1200
Db	1141	GACCAATATGAAGAGATTTTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCATGT	1200
Qy	1201	ACGTGGACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGGTCTTATACGGATAC	1260
Db	1201	ACGTGGACCTTCTCTCGAAATCATTTCTTGTGAGCAAAAGGGTCTTATACGGATAC	1260
Qy	1261	AGCATATCCAAGTTTGTCAATCATTAAGCACACGAGGAGATATATATTTCCATGCGAGAA	1320
Db	1261	AGCATATCCAAGTTTGTCAATCATTAAGCACACGAGGAGATATATATTTCCATGCGAGAA	1320
Qy	1321	AATGATGATGCCAATTTTACCAGGATGTTTACGCTGAATATTAAGAGGAAACCTCAAGTG	1380
Db	1321	AATGATGATGCCAATTTTACCAGGATGTTTACGCTGAATATTAAGAGGAAACCTCAAGTG	1380
Qy	1381	CTCGCAGAGCATCGGCAAGTTCAGCGGCTCTGTTCTCGGATGGATACCCATTTACCATCT	1440
Db	1381	CTCGCAGAGCATCGGCAAGTTCAGCGGCTCTGTTCTCGGATGGATACCCATTTACCATCT	1440
Qy	1441	TGGACCTCGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAGGA	1500
Db	1441	TGGACCTCGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAGGA	1500
Qy	1501	GTCTGGAATAGAAAGGCTTAACAGAAAAGTGTGACAGTGGGTGTCGAGCAGTACTCTA	1560
Db	1501	GTCTGGAATAGAAAGGCTTAACAGAAAAGTGTGACAGTGGGTGTCGAGCAGTACTCTA	1560
Qy	1561	AACATGAGTGAAGCCATTAAGGGTTCCTGCTCAAGTGTGCTGCATACAAATTCCTTGGC	1620
Db	1561	AACATGAGTGAAGCCATTAAGGGTTCCTGCTCAAGTGTGCTGCATACAAATTCCTTGGC	1620
Qy	1621	ACATCTCTGAGACCATCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAAGACAAC	1680
Db	1621	ACATCTCTGAGACCATCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAAGACAAC	1680
Qy	1681	ATCTCATTTCTATGCAACAAATTTGGTGTGTTCTCTCTTTCATTTGCTTTTAAACCTGCTA	1740
Db	1681	ATCTCATTTCTATGCAACAAATTTGGTGTGTTCTCTCTTTCATTTGCTTTTAAACCTGCTA	1740
Qy	1741	ATTTGTCACAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG	1800
Db	1741	ATTTGTCACAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG	1800
Qy	1801	ACCGCTCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATCAATATGATCTC	1860
Db	1801	ACCGCTCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATCAATATGATCTC	1860
Qy	1861	AAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCTTT	1920
Db	1861	AAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAGGTACTAGGATCAGGTGCTTT	1920
Qy	1921	GGAAAGTGTGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGTT	1980
Db	1921	GGAAAGTGTGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGTT	1980
Qy	1981	CCGCTCAAAATGCTGAAGAAAACAGACAGCTCTGAAGAGAGGCACTCATGTGAGAA	2040
Db	1981	CCGCTCAAAATGCTGAAGAAAACAGACAGCTCTGAAGAGAGGCACTCATGTGAGAA	2040
Qy	2041	CTAAGATGATGACCCAGCTGGGAAGCCACAGAAATTTGTGAACCTGCTGGGGGCTGC	2100
Db	2041	CTAAGATGATGACCCAGCTGGGAAGCCACAGAAATTTGTGAACCTGCTGGGGGCTGC	2100
Qy	2101	ACACTGTGAGGACCAATTTACTTTGATTTTGAATACTGTGCTATGGTATCTTCTCAAC	2160
Db	2101	ACACTGTGAGGACCAATTTACTTTGATTTTGAATACTGTGCTATGGTATCTTCTCAAC	2160



```
QY 2161 TATCTAAGAACTAAAGAGAAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
DB 2161 TATCTAAGAACTAAAGAGAAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220
QY 2221 AATTTAGTTTTTACCCCACTTTTCCAAATCACATCAAAATTCAGCATGCCCTGGTTCAAGA 2280
DB 2221 AATTTAGTTTTTACCCCACTTTTCCAAATCACATCAAAATTCAGCATGCCCTGGTTCAAGA 2280
QY 2281 GAAGTTCAAGTACACCCGAGCTCGATCAAAATCTCAGGCTTCATGGGAATTCATTTCCAC 2340
DB 2281 GAAGTTCAAGTACACCCGAGCTCGATCAAAATCTCAGGCTTCATGGGAATTCATTTCCAC 2340
QY 2341 TCTGAAGATGAATTTGAATATGAACCAAAAGGCTGGAAGAGAGAGGAGACTTGAAT 2400
DB 2341 TCTGAAGATGAATTTGAATATGAACCAAAAGGCTGGAAGAGAGAGGAGACTTGAAT 2400
QY 2401 GTGCTTACATTTGAAGATCTCTTTGCTTTTGCATATCAAGTTGCCAAAGGAATGAATTT 2460
DB 2401 GTGCTTACATTTGAAGATCTCTTTGCTTTTGCATATCAAGTTGCCAAAGGAATGAATTT 2460
QY 2461 CTGGAATTTAAGTCTGTGTTTCACAGAGACCTGGCCGCGAGAACTGCTTCTCACCAC 2520
DB 2461 CTGGAATTTAAGTCTGTGTTTCACAGAGACCTGGCCGCGAGAACTGCTTCTCACCAC 2520
QY 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTCAATCCAC 2580
DB 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTCAATCCAC 2580
QY 2581 TATGTTGTACAGGGCAATGCCCGTCTGCTGTAAAATGATGGCCCCCGAAGCCCTGTTT 2640
DB 2581 TATGTTGTACAGGGCAATGCCCGTCTGCTGTAAAATGATGGCCCCCGAAGCCCTGTTT 2640
QY 2641 GAAGGCATCTACACCATTAAGAGTATGCTGGTATATGGAATATTACTGTGGGAATC 2700
DB 2641 GAAGGCATCTACACCATTAAGAGTATGCTGGTATATGGAATATTACTGTGGGAATC 2700
QY 2701 TTCTCACTTGGTGAATCCCTTACCTTGGCATTCGGTTTGTGCTAACTTCTACAACTG 2760
DB 2701 TTCTCACTTGGTGAATCCCTTACCTTGGCATTCGGTTTGTGCTAACTTCTACAACTG 2760
QY 2761 ATTCAAAATGGAATTTAAATGATCAGGCATTTTATGCTACAGAGAATATACATTTATA 2820
DB 2761 ATTCAAAATGGAATTTAAATGATCAGGCATTTTATGCTACAGAGAATATACATTTATA 2820
QY 2821 ATGCAATCTCTGGGCTTTTACCTCAAGAAACGCCATCTTCCCTAAATTTGACTTCG 2880
DB 2821 ATGCAATCTCTGGGCTTTTACCTCAAGAAACGCCATCTTCCCTAAATTTGACTTCG 2880
QY 2881 TTTTATAGGATGTCAGTGGCAGATGCAGAAAGCGATGTATCAGAAATGTGGATGCCGT 2940
DB 2881 TTTTATAGGATGTCAGTGGCAGATGCAGAAAGCGATGTATCAGAAATGTGGATGCCGT 2940
QY 2941 GTTTCGGAATGTCCTCACACCTTACCACAAACAGCGACCTTTTCAGCAGAGAGATGATTTG 3000
DB 2941 GTTTCGGAATGTCCTCACACCTTACCACAAACAGCGACCTTTTCAGCAGAGAGATGATTTG 3000
QY 3001 GGGCTACTCTCCGAGCTCAGTCTGAAGATTGCTAGAGAAACAATTTAGTTTAAAG 3060
DB 3001 GGGCTACTCTCCGAGCTCAGTCTGAAGATTGCTAGAGAAACAATTTAGTTTAAAG 3060
QY 3061 ACTTCATCCCTCCACCTATCCCTACAGGCTGTAGATTACCAAAACAAGATTAATTTCA 3120
DB 3061 ACTTCATCCCTCCACCTATCCCTACAGGCTGTAGATTACCAAAACAAGATTAATTTCA 3120
QY 3121 CACTAAAAGAAAATCTATTATCAACTGCTGCTTCCACAGACTTTTCTCTAGAGCGCGTCT 3180
DB 3121 CACTAAAAGAAAATCTATTATCAACTGCTGCTTCCACAGACTTTTCTCTAGAGCGCGTCT 3180
QY 3181 GCGTTTACTCTGTTTTCAAAGGAGCTTTTGTAAAATCAAAATCATCTCTCACAAGGAG 3240
DB 3181 GCGTTTACTCTGTTTTCAAAGGAGCTTTTGTAAAATCAAAATCATCTCTCACAAGGAG 3240
QY 3241 GAGGAGCTGATTAATGAATTTTATGGAGCATTTGATCTGATCCAGGCCCTTCTCAGGCCG 3300
```

```
DB 3241 GAGGAGCTGATTAATGAATTTTATGGAGCATTTGATCTGATCCCAAGGCTTCTCAGGCCG 3300
QY 3301 CTTTGCAGTGAATTTGTTACCTGAAGTACAGTATATTTCTGTAATACATAAAACAAGAC 3360
DB 3301 CTTTGCAGTGAATTTGTTACCTGAAGTACAGTATATTTCTGTAATACATAAAACAAGAC 3360
QY 3361 ATTTTGCCTAAGGAAGCTAATATATATTTTAAAGTCTATGTTTAAATAATATGATAA 3420
DB 3361 ATTTTGCCTAAGGAAGCTAATATATATTTTAAAGTCTATGTTTAAATAATATGATAA 3420
QY 3421 TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAATTTCTACTACAGA 3476
DB 3421 TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAATTTCTACTACAGA 3476

RESULT 8
HSLT3RTK HSLT3RTK 3422 bp mRNA linear PRI 17-JAN-1995
LOCUS H.sapiens FLT3 mRNA for FLT3 receptor tyrosine kinase.
DEFINITION Z26652 S64785
ACCESSION Z26652.1 GI:406322
VERSION FLT3 gene; receptor tyrosine kinase.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3422)
AUTHORS Rosnet.O., Schiff,C., Pebusque,M.J., Marchetto,S., Tonnelie,C.,
Toiron,Y., Birg,F. and Birnbaum,D.
TITLE Human FLT3/FLK2 gene: cDNA cloning and expression in hematopoietic
cells
JOURNAL Blood 82 (4), 1110-1119 (1993)
MEDLINE 93357464
PUBMED 8394751
REFERENCE 2 (bases 1 to 3422)
AUTHORS Rosnet,O.
TITLE Direct Submission
JOURNAL Submitted (04-OCT-1993) Rosnet O., Inserm U119, 27, bd. Lei Roure.,
Marseille, France
COMMENT data in S64785.
FEATURES
Location/Qualifiers
1..3422
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..2982
/gene="FLT3"
1..2982
/gene="FLT3"
/codon_start=1
/product="FLT3 receptor tyrosine kinase"
/protein_id="CAA81393.1"
/db_xref="GI:406323"
/translat="MPALARDGQPLLVVFSAMIFGTTINODLPVICKVLINHNND
SSVGKSSVPMVSPEDLGCALRQSSGVYRAAEVDSASITLQVLVDAPGNIS
CLWPKHSLNCQPHEDLQNRGVSMVILKMTETQAGEYLLFIQSEATNTLTFTVSI
RNTLYTLRRPIFRKMEQDALVCISEPPIVWVLDGSGESCKEESPAVVKKEE
VHLHFLGMDIRCCARNEQRECTRLFTIDLNOTPOTTLQFLKVGEEPLVCKRAVH
VNHGFLWELENKALEGNYFEMSTYTNTRMIRILFAFVSVAANDTGYTCSK
HPSQALVTIVEKGFNATNSSEDEYIDQEEFCFVRKAYQIRCTWTFSRKSPC
EOKGLDNGYSISKFCNHRHOPGEYIFHAENDDAQTKMTLNIIRKPOVLAEASQA
SCFSDGYPLPSWTWKCKSDKSPNCFEITTEGVNKRANKVFCQWVSSSTLNMSAIK
GELVKCCNSIGTSCETILLNSPGFPFIQNIISFYATIGVCLLFIIVLTLICHKY
KQOFYEQSLQWQVTVSSNDYFYVDFREYEDLWKEFPRENLEFGKVLGSGAFKV
MNATYGISGTVSIQVAMKLEKADSEREALMSLMTQLGSHENIVNLIGACT
LSGPIYLIPEYCCYGLDLNLRSKRKHRTWTEIFEHNFSEYFTFQSHPNSEMPGS
REVQHPDSDIISGLHNSFHEDETEYENOKRLEEDNLVLTFFDLILCFAYQVAKG
MEFLFKSCVHRDLAARNVLYTHGKVKVVICDEGLARDIMSDSNVVRGNARLPVKMA
PSLEFLGTYTIKSDVMSYIGLLWEIFSLGVNYPYGPVDPANFYKLLQNGFKMDQOPYA
TEEYITIMQSWAFDSRKRPSPNLTSLFGCLADAEAEANVQNDVSECPHYONR
RFPFSREMDLGLLSPQAQVEDS"
1..69
sig_peptide
```

mat_peptide		/gene="FLT3" 70..1980		/gene="FLT3" 2983..3422		/product="FLT3 receptor tyrosine kinase"	
3'UTR		polyA_signal 3401..3405		polyA_signal 3401..3405		polyA_signal 3401..3405	
BASE COUNT		1038 a 687 c 758 g 939 t		1038 a 687 c 758 g 939 t		1038 a 687 c 758 g 939 t	
ORIGIN							
Query Match		96.9%		Score 3393.8; DB 9; Length 3422;		Score 3393.8; DB 9; Length 3422;	
Best Local Similarity		99.6%		Pred. No. 0;		Pred. No. 0;	
Matches 3401;		Conservative 0;		Mismatches 12;		Mismatches 12;	
		Indels 0;		Indels 0;		Indels 0;	
		Gaps 0;		Gaps 0;		Gaps 0;	
Qy	58	ATGCCGCGGTTGGCGCGGACCGCGGGACCGTGGCGCTGCTCGTGTGTTTTTCTGCAATG	117	ATGCCGCGGTTGGCGCGGACCGCGGGACCGTGGCGCTGCTCGTGTGTTTTTCTGCAATG	60	ATGCCGCGGTTGGCGCGGACCGCGGGACCGTGGCGCTGCTCGTGTGTTTTTCTGCAATG	177
Db	1	ATGCCGCGGTTGGCGCGGACCGCGGGACCGTGGCGCTGCTCGTGTGTTTTTCTGCAATG	60	ATGCCGCGGTTGGCGCGGACCGCGGGACCGTGGCGCTGCTCGTGTGTTTTTCTGCAATG	177	ATGCCGCGGTTGGCGCGGACCGCGGGACCGTGGCGCTGCTCGTGTGTTTTTCTGCAATG	177
Qy	118	ATATTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTTTAATCAATCAT	177	ATATTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTTTAATCAATCAT	177	ATATTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTTTAATCAATCAT	177
Db	61	ATATTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTTTAATCAATCAT	120	ATATTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTTTAATCAATCAT	120	ATATTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTTTAATCAATCAT	120
Qy	178	AAGAACATGATTCATCAGTGGGGAAGTCATCATCATATCCATGCTATCAGAAATCCCG	237	AAGAACATGATTCATCAGTGGGGAAGTCATCATCATATCCATGCTATCAGAAATCCCG	237	AAGAACATGATTCATCAGTGGGGAAGTCATCATCATATCCATGCTATCAGAAATCCCG	237
Db	121	AAGAACATGATTCATCAGTGGGGAAGTCATCATCATATCCATGCTATCAGAAATCCCG	180	AAGAACATGATTCATCAGTGGGGAAGTCATCATCATATCCATGCTATCAGAAATCCCG	180	AAGAACATGATTCATCAGTGGGGAAGTCATCATCATATCCATGCTATCAGAAATCCCG	180
Qy	238	GAAGACCTCGGTTGGCGTTCAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCGT	297	GAAGACCTCGGTTGGCGTTCAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCGT	297	GAAGACCTCGGTTGGCGTTCAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCGT	297
Db	181	GAAGACCTCGGTTGGCGTTCAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCGT	240	GAAGACCTCGGTTGGCGTTCAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCGT	240	GAAGACCTCGGTTGGCGTTCAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCGT	240
Qy	298	GTGGAAGTGGATGATCTGCTTCCATCACACTGCAAGTGGTGCATGCCCGACGGGAAC	357	GTGGAAGTGGATGATCTGCTTCCATCACACTGCAAGTGGTGCATGCCCGACGGGAAC	357	GTGGAAGTGGATGATCTGCTTCCATCACACTGCAAGTGGTGCATGCCCGACGGGAAC	357
Db	241	GTGGAAGTGGATGATCTGCTTCCATCACACTGCAAGTGGTGCATGCCCGACGGGAAC	300	GTGGAAGTGGATGATCTGCTTCCATCACACTGCAAGTGGTGCATGCCCGACGGGAAC	300	GTGGAAGTGGATGATCTGCTTCCATCACACTGCAAGTGGTGCATGCCCGACGGGAAC	300
Qy	358	ATTTCTCTCTCGGGTCTTAAAGCACAGTCCCTGGAATGCCAGCCACATTTTGATTTA	417	ATTTCTCTCTCGGGTCTTAAAGCACAGTCCCTGGAATGCCAGCCACATTTTGATTTA	417	ATTTCTCTCTCGGGTCTTAAAGCACAGTCCCTGGAATGCCAGCCACATTTTGATTTA	417
Db	301	ATTTCTCTCTCGGGTCTTAAAGCACAGTCCCTGGAATGCCAGCCACATTTTGATTTA	360	ATTTCTCTCTCGGGTCTTAAAGCACAGTCCCTGGAATGCCAGCCACATTTTGATTTA	360	ATTTCTCTCTCGGGTCTTAAAGCACAGTCCCTGGAATGCCAGCCACATTTTGATTTA	360
Qy	418	CAAAACAGAGGAGTGTTCCTCATGTTTGAATAATCAGAGAACCCAGCTCGAGAA	477	CAAAACAGAGGAGTGTTCCTCATGTTTGAATAATCAGAGAACCCAGCTCGAGAA	477	CAAAACAGAGGAGTGTTCCTCATGTTTGAATAATCAGAGAACCCAGCTCGAGAA	477
Db	361	CAAAACAGAGGAGTGTTCCTCATGTTTGAATAATCAGAGAACCCAGCTCGAGAA	420	CAAAACAGAGGAGTGTTCCTCATGTTTGAATAATCAGAGAACCCAGCTCGAGAA	420	CAAAACAGAGGAGTGTTCCTCATGTTTGAATAATCAGAGAACCCAGCTCGAGAA	420
Qy	478	TACCTACTTTTTATTCAGAGTGAAGCTACCAATTACACAATATTTTACAGTGAATTA	537	TACCTACTTTTTATTCAGAGTGAAGCTACCAATTACACAATATTTTACAGTGAATTA	537	TACCTACTTTTTATTCAGAGTGAAGCTACCAATTACACAATATTTTACAGTGAATTA	537
Db	421	TACCTACTTTTTATTCAGAGTGAAGCTACCAATTACACAATATTTTACAGTGAATTA	480	TACCTACTTTTTATTCAGAGTGAAGCTACCAATTACACAATATTTTACAGTGAATTA	480	TACCTACTTTTTATTCAGAGTGAAGCTACCAATTACACAATATTTTACAGTGAATTA	480
Qy	538	AGAAATACCCCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGGAAAACAGAGC	597	AGAAATACCCCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGGAAAACAGAGC	597	AGAAATACCCCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGGAAAACAGAGC	597
Db	481	AGAAATACCCCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGGAAAACAGAGC	540	AGAAATACCCCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGGAAAACAGAGC	540	AGAAATACCCCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGGAAAACAGAGC	540
Qy	598	GCCCTGGTCTCATATCTGAGAGCGTTCCAGAGCGATCTGGGAATGGTGTCTTTGCGAT	657	GCCCTGGTCTCATATCTGAGAGCGTTCCAGAGCGATCTGGGAATGGTGTCTTTGCGAT	657	GCCCTGGTCTCATATCTGAGAGCGTTCCAGAGCGATCTGGGAATGGTGTCTTTGCGAT	657
Db	541	GCCCTGGTCTCATATCTGAGAGCGTTCCAGAGCGATCTGGGAATGGTGTCTTTGCGAT	600	GCCCTGGTCTCATATCTGAGAGCGTTCCAGAGCGATCTGGGAATGGTGTCTTTGCGAT	600	GCCCTGGTCTCATATCTGAGAGCGTTCCAGAGCGATCTGGGAATGGTGTCTTTGCGAT	600
Qy	658	TCACAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAAAGGAGAAAAGTG	717	TCACAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAAAGGAGAAAAGTG	717	TCACAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAAAGGAGAAAAGTG	717
Db	601	TCACAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAAAGGAGAAAAGTG	660	TCACAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAAAGGAGAAAAGTG	660	TCACAGGGGGAAGCTGTAAAGAGAAAGTCCAGCTGTGTTTAAAAAGGAGAAAAGTG	660
Qy	718	CTTCATGAATTTATTTGGGACGGACATAGAGTGTCTGTGCCAGAAATGAACCTGGCAGGNA	777	CTTCATGAATTTATTTGGGACGGACATAGAGTGTCTGTGCCAGAAATGAACCTGGCAGGNA	777	CTTCATGAATTTATTTGGGACGGACATAGAGTGTCTGTGCCAGAAATGAACCTGGCAGGNA	777
Db	661	CTTCATGAATTTATTTGGGACGGACATAGAGTGTCTGTGCCAGAAATGAACCTGGCAGGNA	720	CTTCATGAATTTATTTGGGACGGACATAGAGTGTCTGTGCCAGAAATGAACCTGGCAGGNA	720	CTTCATGAATTTATTTGGGACGGACATAGAGTGTCTGTGCCAGAAATGAACCTGGC	





Db 361 CAAACAGAGAGTGTGTTTCCATGGTCATTTTGAAATGACAGAAACCAAGCTGGAGAA 420  
Qy 478 TACCTACTTTTATTACAGAGTGAAGCTACCAATACACAATATTGTTTACAGTGAGTATA 537  
Db 421 TACCTACTTTTATTACAGAGTGAAGCTACCAATACACAATATTGTTTACAGTGAGTATA 480  
Qy 538 AGAANTACCTGCTTTACACATTAAGAAGACCTTACCTTTAGAAAATGGAACCAAGGAC 597  
Db 481 AGAANTACCTGCTTTACACATTAAGAAGACCTTACCTTTAGAAAATGGAACCAAGGAC 540  
Qy 598 GCCCTGCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGGATGGGTGCTTTGGCAT 657  
Db 541 GCCCTGCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGGATGGGTGCTTTGGCAT 600  
Qy 658 TCACAGGGGAAAGCTGTAAAGAAAGAGTCCAGCTGTGTTTAAAGAGGAGGAAAAAGTG 717  
Db 601 TCACAGGGGAAAGCTGTAAAGAAAGAGTCCAGCTGTGTTTAAAGAGGAGGAAAAAGTG 660  
Qy 718 CTTTCATGAATATTGTTGGAGCGGACATAAAGTGTGTTGCCAGAAATGAACCTGGGAGGAA 777  
Db 661 CTTTCATGAATATTGTTGGAGCGGACATAAGTGTGTTGCCAGAAATGAACCTGGGAGGAA 720  
Qy 778 TGCACCAAGCTGTTCAACAATAGATCTAAATCAAACTCCTCAGACACCAATTGCCACAATTA 837  
Db 721 TGCACCAAGCTGTTCAACAATAGATCTAAATCAAACTCCTCAGACCAATTTGCCACAATTA 780  
Qy 838 TTTCTTAAAGTAGGGGAACCTTATGGATAAGGTGCAAGAGCTGTTCATGTCAACCATGGA 897  
Db 781 TTTCTTAAAGTAGGGGAACCTTATGGATAAGGTGCAAGAGCTGTTCATGTCAACCATGGA 840  
Qy 898 TTCGGGCTACCTCGGGAATTAGAAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATG 957  
Db 841 TTCGGGCTACCTCGGGAATTAGAAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATG 900  
Qy 958 AGTACCTATTCAACAACAGAACTATGATACGGATTCGTGTTGCTTTTGTATCATCAGTG 1017  
Db 901 AGTACCTATTCAACAACAGAACTATGATACGGATTCGTGTTGCTTTTGTATCATCAGTG 960  
Qy 1018 GCAAGAAACGACACCGGATCTACACTTGTTCCTTTCAAGAGCTCCAGTCAATCAGCT 1077  
Db 961 GCAAGAAACGACACCGGATCTACACTTGTTCCTTTCAAGAGCTCCAGTCAATCAGCT 1020  
Qy 1078 TTGGTTACCATCGTAGGAAAGGGATTTATAATGCTACCAATTTCAAGTGAAGATATTGAA 1137  
Db 1021 TTGGTTACCATCGTAGGAAAGGGATTTATAATGCTACCAATTTCAAGTGAAGATATTGAA 1080  
Qy 1138 ATTGACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCTACCCACAAATCAGA 1197  
Db 1081 ATTGACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGGCTACCCACAAATCAGA 1140  
Qy 1198 TGTACGTGGACCTTCTCTCGAAAATCATTTCTTGTGAGCAAAAGGGTCTTTGATAACGGA 1257  
Db 1141 TGTACGTGGACCTTCTCTCGAAAATCATTTCTTGTGAGCAAAAGGGTCTTTGATAACGGA 1200  
Qy 1258 TACAGCATATCCAAAGTTTTCGAATCATAAAGCACCGAGGAGAGATATATATTCATGCA 1317  
Db 1201 TACAGCATATCCAAAGTTTTCGAATCATAAAGCACCGAGGAGAGATATATATTCATGCA 1260  
Qy 1318 GAAANTGATGATGCCCAATTTACCAAAATGTTACGCTGGAATATAAGAAGGAAACCTCAA 1377  
Db 1261 GAAANTGATGATGCCCAATTTACCAAAATGTTACGCTGGAATATAAGAAGGAAACCTCAA 1320  
Qy 1378 GTGCTCCGAGAGCATCGGCAAGTTCAGCGTCTGTTTCTCGGATGGATACCCATTACCA 1437  
Db 1321 GTGCTCCGAGAGCATTCGGCAAGTTCAGCGTCTGTTTCTCGGATGGATACCCATTACCA 1380  
Qy 1438 TCTTGGACCTCGGAAGAAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAA 1497  
Db 1381 TCTTGGACCTCGGAAGAAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAA 1440  
Qy 1498 GGAGTCTGGAAATAGAAAGGCTTAACAGAAAAGTGTGTTGACAGAGTGGGTGTCGAGCAGTACT 1557  
Db 1441 GGAGTCTGGAAATAGAAAGGCTTAACAGAAAAGTGTGTTGACAGAGTGGGTGTCGAGCAGTACT 1500

Qy 1558 CTAACATGAGTGAAGCCATAAAAGGGTTCCTGGTCAAGTGTGTGCATACAAATTCCTTT 1617  
Db 1501 CTAACATGAGTGAAGCCATAAAAGGGTTCCTGGTCAAGTGTGTGCATACAAATTCCTTT 1560  
Qy 1618 GGCACATCTTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCTCAAGAC 1677  
Db 1561 GGCACATCTTGTGAGACGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCTCAAGAC 1620  
Qy 1678 AACATCTCATCTATGCAACAAATTTAGTGTCTCTCCCTTTCATCTGCTCTTTTAAACCTTG 1737  
Db 1621 AACATCTCATCTATGCAACAAATTTAGTGTCTCTCCCTTTCATCTGCTCTTTTAAACCTTG 1680  
Qy 1738 CTAATTTGTCACAGTACAAAAGCAATTTAGTGTGAAAGCCAGCTACAGATGGTACAG 1797  
Db 1681 CTAATTTGTCACAGTACAAAAGCAATTTAGTGTGAAAGCCAGCTACAGATGGTACAG 1740  
Qy 1798 GTGACCGGCTCCTCAGATAATAGAGTACTTACGTTGATTTTACAGAGATATGAATATGAT 1857  
Db 1741 GTGACCGGCTCCTCAGATAATAGAGTACTTACGTTGATTTTACAGAGATATGAATATGAT 1800  
Qy 1858 CTCRAATCGGAGTTCCTCAAGAGAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCT 1917  
Db 1801 GTCAAAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCT 1860  
Qy 1918 TTTGAAAAGTGCATCAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG 1977  
Db 1861 TTTGAAAAGTGCATCAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG 1920  
Qy 1978 GTTCCGCTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGACACTCATGTCA 2037  
Db 1921 GTTACCCTGCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGACACTCATGTCA 1980  
Qy 2038 GAACTCAGATGATGACCCAGCTGGGAAGCCAGAGATATTTGTGAACCTGCTGGGGGGG 2097  
Db 1981 GAACTCAGATGATGACCCAGCTGGGAAGCCAGAGATATTTGTGAACCTGCTGGGGGGG 2040  
Qy 2098 TGCACACTGTGAGGACCAATTTACTTTGATTTTGAATACTGTTGCTATGCTGATCTCTTC 2157  
Db 2041 TGCACACTGTGAGGACCAATTTACTTTGATTTTGAATACTGTTGCTATGCTGATCTCTTC 2100  
Qy 2158 AACTATCTAAGAAGTAAAGGAAAATTTTCAGGACTTGGACAGAGATTTTTCAAAGGAA 2217  
Db 2101 AACTATCTAAGAAGTAAAGGAAAATTTTCAGGACTTGGACAGAGATTTTTCAAAGGAA 2160  
Qy 2218 CACAAATTTTACCCACTTTTCCAAATCACAATCCAAATTTCCAGCATGCTGCTTCA 2277  
Db 2161 CACAAATTTTACCCACTTTTCCAAATCACAATCCAAATTTCCAGCATGCTGCTTCA 2220  
Qy 2278 AGAGAAGTTCAGATACACCGGACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTT 2337  
Db 2221 AGAGAAGTTCAGATACACCGGACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTT 2280  
Qy 2338 CACTCTGAAGATGAAATTAATATGAAACCAAAAAGGCTGGAAGAGAGGAGGACTTG 2397  
Db 2281 CACTCTGAAGATGAAATTAATATGAAACCAAAAAGGCTGGAAGAGAGGAGGACTTG 2340  
Qy 2398 AATGCTCTACATTTGAAGATCTTCTTGTGCTTTGATATCAAGTTGCCAAGGAATGGAA 2457  
Db 2341 AATGCTCTACATTTGAAGATCTTCTTGTGCTTTGATATCAAGTTGCCAAGGAATGGAA 2400  
Qy 2458 TTTCTGGAATTTAAGTCTGCTTCCAGAGACCTTGCCCGCCAGGAACTGCTTGTGCACC 2517  
Db 2401 TTTCTGGAATTTAAGTCTGCTTCCAGAGACCTTGCCCGCCAGGAACTGCTTGTGCACC 2460  
Qy 2518 CACGGAAAGTGGTGAAGATATGTGACTTTTGGATTTGGCTCGAGATATCATGAGTATTC 2577  
Db 2461 CACGGAAAGTGGTGAAGATATGTGACTTTTGGATTTGGCTCGAGATATCATGAGTATTC 2520  
Qy 2578 AACTATGTTGTCAGGGCAATCCCGCTGCTGCTGTAATAATGATGCTCCCGCCAGGAACTG 2637  
Db 2521 AACTATGTTGTCAGGGCAATCCCGCTGCTGCTGTAATAATGATGCTCCCGCCAGGAACTG 2580

```
QY 2638 TTTGAAGGCTCTACACCATTAAGAGTGATGCTGTGGTCATATGGAATATTACTGTGGAA 2697
Db 2581 TTTGAAGGCTCTACACCATTAAGAGTGATGCTGTGGTCATATGGAATATTACTGTGGAA 2640
QY 2698 ATCTTCTCACTTGCTGTGAATCCTTACCCCGGCAATCCGGTGTAGTGTAACCTTCTACAAA 2757
Db 2641 ATCTTCTCACTTGCTGTGAATCCTTACCCCGGCAATCCGGTGTAGTGTAACCTTCTACAAA 2700
QY 2758 CTGATTCAAAATGGAATTAATAATGGATCAGCCATTTTATGCTACAGAGAAATATACATT 2817
Db 2701 CTGATTCAAAATGGAATTAATAATGGATCAGCCATTTTATGCTACAGAGAAATATACATT 2760
QY 2818 ATATGAATCCTCTGGGCTTTTGAATCAAGGAAAGCGGCATCTCCCTTAATTTGACT 2877
Db 2761 ATATGAATCCTCTGGGCTTTTGAATCAAGGAAAGCGGCATCTCCCTTAATTTGACT 2820
QY 2878 TCGTTTTTAGGATGTCAGCTGGCAGATGCGAAGAAGCGATGTATCAGAATGTGGATGGC 2937
Db 2821 TCGTTTTTAGGATGTCAGCTGGCAGATGCGAAGAAGCGATGTATCAGAATGTGGATGGC 2880
QY 2938 CGTGTTCGGAATGTCCTCACACCTACCAAAACAGGCGACCTTTTCAGCAGAGATGGAT 2997
Db 2881 CGTGTTCGGAATGTCCTCACACCTACCAAAACAGGCGACCTTTTCAGCAGAGATGGAT 2940
QY 2998 TTGGGGCTACTCTCCGCGAGCTCAGGTGGAAGATTCGTAGAGAAACAATTTAGTTTTTA 3057
Db 2941 TTGGGGCTACTCTCCGCGAGCTCAGGTGGAAGATTCGTAGAGAAACAATTTAGTTTTTA 3000
QY 3058 AGGACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAAGATTAATTT 3117
Db 3001 AGGACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAAGTAAATTT 3060
QY 3118 CATCACTAAAAGAAAATCTATATCACTGCTGTCTTACCAGACTTTTCTCTAGAGCGG 3177
Db 3061 CATCACTAAAAGAAAATCTATATCACTGCTGTCTTACCAGACTTTTCTCTAGAGAGCG 3120

RESULT 10
LOCUS I44732 3521 bp DNA linear PAT 07-OCT-1997
DEFINITION Sequence 1 from patent US 5635388.
ACCESSION I44732
VERSION I44732.1 GI:2469445
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3521)
AUTHORS Bennett,B.D., Broz,S.D., Matthews,W. and Zeigler,F.C.
TITLE Agonist antibodies against the flk2/flt3 receptor and uses thereof
JOURNAL Patent: US 5635388-A 1 03-JUN-1997;
FEATURES
Location/Qualifiers
Source
BASE COUNT 943 a 854 c 907 g 817 t
ORIGIN
Query Match 65.0%; Score 2274.6; DB 6; Length 3521;
Blast Local Similarity 80.3%; Pred. No. 0;
Matches 2805; Conservative 0; Mismatches 649; Indels 41; Gaps 10;

QY 17 GCGTGGCGGGCGCCCTGGGGGACCCCGGGCTCCGGAGGCCATGCCGGCTTGGC---GC 73
Db 42 GCGTGGCGGGCGGGCTTGGGGGACCCCGGGCTCCGGAGGCCATGCCGGCTTGGC---GC 100
QY 74 GCGAGCGGGGACCGGTGGCGCTGCTGTTGTTTTCTGCAATGATATTGGGACTATTTA 133
Db 101 GCGAGCGGGGCGGGCTGCTGCTGCTGTTGTTTGTCAAGTAATGATCTTGGACCGTTA 160
QY 134 CAAATCAAGATGCTGCTGATCAAGTGTTTAAATCAATCAATCAATCAATCAATCAATCAAT 193
Db 161 CAAATCAAGATGCTGCTGATCAAGTGTTTAAATCAATCAATCAATCAATCAATCAATCAAT 220
```

```
QY 194 CAGTGGGGAAGTCATCATCATATATCCATGTGATCAGAAATCCCGGGAACACCTCGGCTGTG 253
Db 221 CAGCGGGAAGCCATCATCGTACCGAATGGTGGAGGATCCCGAGAACACCTCCAGTGA 280
QY 254 CGTTGAGACCCACAGCTCAGGACAGTGACGAGCTGCCGCTGTGGAAGTGGATGTAT 313
Db 281 CCCGAGGCGCGCAGAGTGAAGGACGGTATATGAAGCGCGCACCTCCCTGAGGTGGCCGAGT 340
QY 314 CTGCTTCATCATCAGCTGCAAGTGTGCTGATGCCCGAGGAAACATTTCTCTCTCTGGG 373
Db 341 CTGGGTCTCATCAGCTGCAAGTGTGAGTGCAGCTGCCACCCAGGGGACCTTTCTCTCTGGG 400
QY 374 TCTTTAAGCAGACGCTCCCTGAATTTGCGAGCCACATTTTGAATTTTCAAAACAGAGAGTTG 433
Db 401 TCTTTAAGCAGACGCTCCCTGGCTGCCAGCGCACCTTTGATTTTCAAAACAGAGAAATCG 460
QY 434 TTTCCATGGTGTATTTGAAATGACAGAAACCCCAAGCTGGAGAAATACCTACTTTTATTC 493
Db 461 TTTCCATGGCTCTTTGAACGTGACAGAGACCCAGGAGGAAATACCTACTTCCATATTTC 520
QY 494 AGAGTGAAGCTTACCAATTTACAAATATTGTTTACAGTGTAGTATTAAGAAATACCTCTTT 553
Db 521 AGAGCGAAGCGCCCACTTACACAGTACTGTTTACAGTGAATGTAAAGATACACAGCTGT 580
QY 554 ACACATTAAGAGACCTTACTTTAGAAAATGGAACACAGAGCGCCCTGTCTGTGATAT 613
Db 581 ATGCTGAAGGAGACCTTACTTTAGGAAGATGGAACACAGGATGCACCTGCTCTGCATCT 640
QY 614 CTGAGAGCGTTCACAGCGGATCGTGGATGGTGGTTCGCGATTCACAGGGGGAAGCT 673
Db 641 CCGAGGGTGTTCGCGAGGCCACTGTGGAGTGGTGTCTGCGAGCTCCACAGGGAAGCT 700
QY 674 GTAAAGAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTGTCTATGATTTTGG 733
Db 701 GTAAAGAGAAAGCGCTGCTGTGTCAGAAAGAGGAAAGGTACTTATGATTTCTTCG 760
QY 734 GGACGACATAAAGTGTGTGCCAGAAATGAATGGGAGGGAAGTGCACAGCGCTGTTC 793
Db 761 GAACAGACATCAGATGCTGTGCTAGAAATGCATGGGCGCGCAATCGACCAAGCTGTTC 820
QY 794 CAATAGATTAATCAACTCCTCAGACCAATGCGCAACAATTTTCTTAAAGTAGGGG 853
Db 821 CCATAGATCAAAACAGGCTCCTCAGAGCACACTGCCAGATTATTTCTTGAAGTGGGG 880
QY 854 AACCTTATGATTAAGTGCAGAGCTGTTTCAAGTGAACCATGATTCGGGCTCACCTGGG 913
Db 881 AACCTTGTGATCAGGTGTAAGGCCATCCATGTGAACCATGATTCGGGCTCACCTGGG 940
QY 914 AATTAGAAAACAAAGACCTCAGGAGGCACTTTTGAGATGAGTACCTATTCAACAA 973
Db 941 AGCTGGAAGACAAAGCCCTGGAGGAGGCGAGCTTCTTGAGATGAGTACCTACTTCCACAA 1000
QY 974 ACAGAACTATGATACGGATTCGTTTCTTATCATCAGTGGGAAGAAACGACACCG 1033
Db 1001 ACAGGACCATGATTCGGATTCCTTGGCTTCTTGGTTCCTCCGTTGGGAAGAAACGACACCG 1060
QY 1034 GATCTACACTTGTCTTCTTCAAGCATCCAGTCAATCAGCTTTGGTTTACCATCTAG 1093
Db 1061 GATATTACACTGCTCTTCTTCAAGCACCCAGCAGCAGCTGTTGGTTCACCATCTAG 1120
QY 1094 GAAAGGATTTTAAATGCTTACCAATTCAGTGAAGTATTAAGTAAATGACCAATATGAAG 1153
Db 1121 AAAAAGGTTTAAAGGCTTACCAGCTCGCAAGAAAGATTAAGTAAATGACCCGCTACGAAA 1180
QY 1154 AGTTTGTCTGTGAGTTTAAAGCTTACCACCAATACAGATGTACCTGGACCTTCT 1213
Db 1181 AGTTTGTCTGTGAGTTTAAAGCTTACCACCAATCCGATCCGATGACCTGGATCTTCT 1240
QY 1214 CTCGAAATCATTTCTTGTGAGCAAAAGGCTCTTGATAACGATACAGATATTCAGT 1273
Db 1241 CTCGAAATCATTTCTTGTGAGCAAGAGGCTTGGAGAGTGGTACAGCATATCTAAT 1300
QY 1274 TTTGCAATCAAGACACCGAGGAGATATATATTTCATGCAGAAATATGATGCC 1333
```

1301 TTTGCGATCATAAGAACAGCCAGGAGAGTACATATTCTATGCGAAAAATGATGCGCCC 1360  
1334 AATTACCAAAATGTTACCGCTGATATTAAGAAAGAAACCTCAAGTGTCCGAGAAAGCAT 1393  
1361 AGTTACCAAAATGTTACCGCTGATATTAAGAAAGAAACCTCAAGTGTCTGCAAAATGCGT 1420  
1394 CGGCAAGTCAGGCGCTCTGTTCTTCGATGATACCCATTACCATTCTTTGGACCTGGAAGA 1453  
1421 CAGCCAGCGCGCTCTGTTCTTCGATGATGCTGATGCTACCCGCTACCTCTTTGGACCTGGAAGA 1480  
1454 AGTGTTCAGACAAGTCTCCCAACTGSCACAGAAAGATCACAGAAGAGTCTGGATAGAA 1513  
1481 AGTGTTCGGAACAAATCTCCCAATTCACAGGAGAAATCCCAAGAGGAGTCTTGGAAATAA 1540  
1514 AGGCTAACAGAAAAGTGTTCGAGAGTGGTGTCCGAGCAGTACTCTAAACATGAGTGAAG 1573  
1541 AGGCTAACAGAAAAGTGTTCGAGAGTGGTGTCCGAGCAGTACTCTAAATATGAGTGAAG 1600  
1574 CCATAAAAGGGTTCCTGTGCTCAAGTGTGTCATACAAATTCCTTGGCACATCTTGTGAGA 1633  
1601 CCGGGAAGGGCTTCTGGTCAAAATGCTGTGCTGATATTCATGGGCACGCTCTTGGCGAAA 1660  
1634 CGATCTCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGACAACATCTCATTTCTATG 1693  
1661 CCATCTTTTAAACTCACAGGCCCTTCCCTTTTCATCCAAAGACAACATCTCTCTTCTATG 1720  
1694 CAACAATTTGTTGTTCT 1753  
1721 CGACCATTTGGGCTCTGCT 1780  
1754 AAAAAAGCAATTTAGTGTATGAAGCCAGCTACAGATGGTACAGTGACGCGTCCCTCAG 1813  
1781 AAAAAAGCAATTTAGTGTACAGAGTACAGTGTGCGATGATCAGTGTGACTGGCCCCCTGG 1840  
1814 ATAATGAGTACTTCTACCTTTGATTCAGAGAAATATGAATATGATCTCAAAATGGAGTTTC 1873  
1841 ATAACGAGTACTTCTACCTTTGATTCAGAGAAATATGAATATGATCTCAAAATGGAGTTC 1900  
1874 CAAGAGAAATTTAGAGTTTGGGAAGGTACTAGATCAGAGTGTCTTTGGAAAAAGTATGA 1933  
1901 CGAGAGAAATTTAGAGTTTGGGAAGGTCTCTGGGCTGTGGGCTTTCGGGAGGTGATGA 1960  
1934 ACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTTGCCGCTCAAAATGC 1993  
1961 ACGCCAGCGCTATGAGCATTTAGTAAACGGGAGTCTCAATTCAGTGTGCGGTGAAGATGC 2020  
1994 TGAAGAAAAAGCAGACAGCTCTGAAAGAGAGGACCTCATGTCAGACATCAAGATGATGA 2053  
2021 TAAAAAGAAAAAGCTGCAGAGCTGTCAAAAAAGAGCTCTCATGTCGAGCTCAAAATGATGA 2080  
2054 CCGAGCTGGGAGCCAGAGAAATTTGTGAACCTGCTGGGGGCGGTGCACATCTGCAGAC 2113  
2081 CCCACCTGGGACACATGACAACATCGTGAATCTGCTGGGGGCAATGCACATCTGCAGGC 2140  
2114 CAATTTACTTTGATTTTGAATACTGTTGCTATGCTGATCTCTCAACTATCTAAGAAGTA 2173  
2141 CAGTGTACTTTGATTTTGAATATTTGTTGCTATGCTGACCTCTCAACTTACCTAAGAAGTA 2200  
2174 AAAGAAAAAATTTACAGGACTTTGGACAGAGATTTTCAAGGAACACAAATTTCAAGTTT 2233  
2201 AAAGAGAGAGTTTACAGGACATGGACAGAGATTTTAAAGACATAAATTTCAAGTTT 2260  
2234 ACCCCTTTTCCAAATCAATCCAAATTTCCAGCATGCTGGTTCAGAGAAAGTTCAGATAC 2293  
2261 ACCCTACTTTCCAGGCACATTCAAATTTCCAGCATGCTGGTTCAGAGAAAGTTCAGTTAC 2320  
2294 ACCGGACTCGGATCAATCTCAGGCTTTCATGGGAATTCATTTTCATCTCGAAGTAA 2353  
2321 ACCCGCCCTTTGGATCAGCTCTCAGGTTTCAATGGGAATTCATTTTCATCTCGAAGTAA 2380  
2354 TTGAATATGAACCAAAAAAGGCTG-----GAAGAGAGGAGGACTTGAATCTGCTTA 2407

2381 TTGAATATGAACCAACCAAGAGGCTGGCAGAAAGAGAGGAGGAATTTGAACGTGCTGA 2440  
2408 CATTTTCAAGATCTTTTTCGCTTTGCATATCAAGTTGCCAAAGAAATTTCTGGAAT 2467  
2441 CGTTTGAAGACCTCTTTTTCGCTTTGCTACCAAGTGGCCAAAGGCATGGAATTTCTCGAGT 2500  
2468 TTAAGTCTGTGTTTTCACAGAGACCTGGCGCCAGGAACTGCTTGTACCCACGGGAAG 2527  
2501 TCAAGTCTGTGTTTTCACAGAGACCTGGCAGCCAGGAATGTGTTGGTACCCACGGGAAG 2560  
2528 TGTGAAAGATATGTCACTTTGGATTGGCTCGAGATATCATGATGATTTCAAACTATGTTG 2587  
2561 TGTGAAAGATCTGTGACTTTGGACTGGCCGAGACATCTGAGCGACTCCAGCTACGTCG 2620  
2588 TCAGGGGCAATGCGCTGTGCTGTAAATGATGGCCGCCCGGAAAGCCCTGTTTGAAGGCA 2647  
2621 TCAGGGGCAACACGCGCTGCGGTGAAGTGGATGGCACCCGAGAGCTTATTTGAAGGGA 2680  
2648 TCTACACCATTAAGAGTGTGCTGTGCTCATATTTGGAATATTACTGTGGGAAATCTTCTCAC 2707  
2681 TCTACACAATCAAGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2740  
2708 TTGTTGTGAATCTTACCCTGGCATTCGCTTACCTTATGCTTACCTTCTACAACTGATTTCAA 2767  
2741 TGGGTGTGAACCTTACCCTGGCATTCCTGTGCTGACGCTAACTTCTATATAACTGATTCAGA 2800  
2768 ATGATTTTAAATGATGATCAGCCATTTTATGCTACAGAAATATATACATTAATGCAAT 2827  
2801 GTGGATTTAAATGAGCAGCCATTTCTATGCCACAGAGGATATATCTTTGTAATGCAAT 2860  
2828 CCGTGTGGCTTTTTCACCAAGAAAGCCATCTTCCCTCAATTTGATGCTGCTGCTGCTGCTGCT 2887  
2861 CCGTGTGGCTTTTTCACCAAGAAAGCCATCTTCCCTCAATTTGATGCTGCTGCTGCTGCTGCT 2920  
2888 GATGTCAGCTGGCAGATGCAGAAAGCGATATATCAGAAATGTGGATGGCCGCTGTTTCGG 2947  
2921 GATGTCAGCTGGCAGAGCAGAAAGCGATATATCAGAAATGTGGATGGCCGCTGTTTCGG 2980  
2948 AATGTCCTCACACTTACCACAAACAGGCGACCTTTTACAGCAGAGATGGAATTTGGGCGTAC 3007  
2981 AATCTCCATCTCATCTACCAAAACAGGCGCCCTCTCAGCAGAGAGGCGGCTCAGAGCCGC 3040  
3008 TCTCTCCAGCGCTCAGGTCGGAAGATTCGTACAGGAACAATTTAGTTTAAAGGACATTCAT 3067  
3041 CATGCCACAGCCCGAGGT-GAAGATTCACAGAAAGAAAGTTAGCGAGAGGCGCTTTGGA 3099  
3068 CCCTCCACCTTATCCCTAACAGCGCTGTAGATTACCAAAACAAGATTAATTTTCATCACTAAA 3127  
3100 CCGCGCCA-----CCCTAGCAGCTGTAGACCCGAGAGCAAGATTAGCCTCGCCTCT--G 3153  
3128 AGAAATCTATTATCACTGTCTTACAGACATTTTCTTAGAAGCGCTGTGCGTTTA 3187  
3154 AGGAAGCGCCCTACAGGCGCTTGTCTGCTGAGCTTTTCTCTAGATGCTGTCTGCGCATTA 3213  
3188 CTCCTGTTTTCAAAGGAGCTTTTGAATCAAAATCAATCATCTCTCACAAGCGGAGGAGGC 3247  
3214 CTC-----CAAGTGACTTCTATAAATCAAACTCTCTCCGACAGGCGGAGAGGC 3266  
3248 TGATAATGAACCTTTATTTGGAGCATTTGCTGTGATCAAGCGCTTCTCAGCGCGGCTTGAG 3307  
3267 AATATGAGACTTGTGTTGAGCCCGCTACCTCGGGGGCGCTTTCCAGGCGCCCCCAGGC 3326  
3308 TGAATG-----TGACTGAAAGTACAGTATATTTCTTGAATATACATAAAACAAA-- 3357  
3327 TTGAGGGAAGCCATGTATCTGAAATATAGTATATCTTGTGAATACGTGGAACAAACC 3386  
3358 -----AGCATTTTGTAGGAGACCTAATATGATTTTT--AGTCTATGTTTTAAAT 3410  
3387 AAACCCGTTTTTGTGTAAGGAAAGCTAAATATGATTTTTTAAAAATCTATGTTTTAAAT 3446  
3411 AATATGAAATTTTTCAGCTATTTAGTATATTTTATGTTGGTGGAAATAAATTTCTAC 3470  
3447 ACTATGTAACCTTTTTCATCTATTTAGTATATTTTATGATGGAATAAATTTCTTAC 3506



Db	1121	AAAAAGGGTTTATAAAGCCTACCAAGCTCGCAAGAGAGTATGAATTAACCCGTACGAAA	1180
Qy	1154	AGTTTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGTACGTGGACCTTCT	1213
Db	1181	AGTTCTGCTCTCACTAGGTTTAAAGCGTACCCACCAATCCGATGACGTGGATCTTCT	1240
Qy	1214	CTCGAAATCATTTCCCTTGTGAGCAAAAGGCTTTGATACGGATACAGCATATCCCAAGT	1273
Db	1241	CTCAAGCCTCATTTCCCTTGTGAACAGAGAGCCCTGGAGGATGGGTACAGCATATCTAAAT	1300
Qy	1274	TTTGCATCATTAAGCACCAGCAGGAGAAATATATATTCATGCAAGAAATGATGATGCC	1333
Db	1301	TTTGCATCATTAAGCAAGCAGGAGAGTATATTTCTATSCAGAAATGATGATGCC	1360
Qy	1334	AAATTACCAAAATGTTTCAAGCTGAATATAGAAGCAACCTCAAGTGTCTCCCAAGAGCAT	1393
Db	1361	AGTTACCAAAATGTTTCAAGCTGAATATAGAAGCAACCTCAAGTGTCTCCCAAGAGCAT	1420
Qy	1394	CGCAAGCTCAGGCGTCTGTTTCTCGATGGATACCCATTAACCATTTGGACCTTGGAGA	1453
Db	1421	CAGCCAGCCAGGCGTCTGTTTCTCGATGGATACCCATTAACCATTTGGACCTTGGAGA	1480
Qy	1454	AGTGTTCAGACAAGTCTCCCACTGCACAGAGATACAGAAAGAGTCTGGAAATAGAA	1513
Db	1481	AGTGTTCAGACAAGTCTCCCACTGCACAGAGATACAGAAAGAGTCTGGAAATAGAA	1540
Qy	1514	AGCTTAACAGAAAGTGTGTCAGAGTGGTGTGAGCAGTACTCTAAACATGAGTGAAG	1573
Db	1541	AGCTTAACAGAAAGTGTGTCAGAGTGGTGTGAGCAGTACTCTAAACATGAGTGAAG	1600
Qy	1574	CCATAAAGGGTCTGCTCAAGTGTGTGCATACAATTCCTTGGCACATCTTTGTGAGA	1633
Db	1601	CCGGAAAGGCTCTGCTCAAGTGTGTGCATACAATTCCTTGGCACATCTTTGTGAGA	1660
Qy	1634	CGATCCTTTTAACTCTCCAGGCCCCCTCCCTTTTCATCCAGACAACATCTCATTTATG	1693
Db	1661	CCATCTTTTAACTCTCCAGGCCCCCTCCCTTTTCATCCAGACAACATCTCATTTATG	1720
Qy	1694	CAACAATTTGCTGTTGCTCTCTTCAATTTGCTGTTTAAACCTGCTCAATTTGTGCACAGT	1753
Db	1721	CGACCAATTTGCTGTTGCTCTCTTCAATTTGCTGTTTAAACCTGCTCAATTTGTGCACAGT	1780
Qy	1754	ACAAAAGCAATTTAGTATGAAGCCAGCTACAGTGTACAGTGTACAGGCTCCCTCAG	1813
Db	1781	ACAAAAGCAATTTAGTATGAAGCCAGCTACAGTGTACAGTGTACAGGCTCCCTCAG	1840
Qy	1814	ATAATGAGTACTTCTAGTGTATTCAGAGAAATATGAATATGATCTCAAAATGGAGTTTC	1873
Db	1841	ATAACGAGTACTTCTAGTGTATTCAGAGAAATATGAATATGATCTCAAAATGGAGTTTC	1900
Qy	1874	CAAGAGAAATTTAGAGTTTGGCAAGGTACTAGATCAGTGTCTTTTGGAAAGATGATGA	1933
Db	1901	CGAGAGAAATTTAGAGTTTGGCAAGGTCTCTGGGCTCTGGGCTCTGGGAGGTTGATGA	1960
Qy	1934	ACGCAACAGCTTTATGAATTTAGCAAAACAGGAGTCTCAATCCAGGTTTGGCGTCAAAATGC	1993
Db	1961	ACGCCAGGCTTATGCAATTTAGTAAACAGGAGTCTCAATCCAGGTTTGGCGTCAAAATGC	2020
Qy	1994	TGAAGAAAGCAGACAGCTCTGAAGAGAGGAGCTCATGTCCAGAACTCAAGATGATGA	2053
Db	2021	TAAAGAGAAAGCTGACAGCTGTGAAAAGCAAGCTCTCATGTCCGAGCTCAAAATGATGA	2080
Qy	2054	CCCAGTGGGAAGCCAGGAGAAATTTGTGAACCTGCTGGGGGGTGCACACTGTCCAGGAC	2113
Db	2081	CCCAGTGGGAAGCCAGGAGAAATTTGTGAACCTGCTGGGGGGTGCACACTGTCCAGGAC	2140
Qy	2114	CAATTTACTTGTATTTTGAATCTGTTGATGATGTTCTCAACTATCTAAGAGTA	2173
Db	2141	CAGTGTACTTGTATTTTGAATCTGTTGATGATGTTCTCAACTATCTAAGAGTA	2200
Qy	2174	AAAGAGAAATTTTCAAGGAGTGTGGACAGAGATTTTCAAGGAGCAACAATTTTCAAGTTT	2233
Db	2201	AAAGAGAGATTTTCAAGGAGTGTGGACAGAGATTTTCAAGGAGCAACAATTTTCAAGTTT	2260
Qy	2234	ACCCACCTTTCCAAATCACATCCAAATTCAGCATGCCCTGGTTCAAGAGAGTTCATCATAC	2293
Db	2261	ACCCACCTTTCCAAATCACATCCAAATTCAGCATGCCCTGGTTCAAGAGAGTTCATCATAC	2320
Qy	2294	ACCCGAGCTCGGATCAAAATCTCAGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAA	2353
Db	2321	ACCCGAGCTCGGATCAAAATCTCAGGCTTCATGGGAATTCATTTCACTCTGAAGATGAAA	2380
Qy	2354	TTGAATATGAAGCAACCAAAAGGCTG-----GAAGAGAGGAGGACTTTGAATGTCTTA	2407
Db	2381	TTGAATATGAAGCAACCAAAAGGCTG-----GAAGAGAGGAGGACTTTGAATGTCTTA	2440
Qy	2408	CATTGGAAGATCTCTTTTGTCTTGTGATATCAAGTTTCCCAAGAGATGGAATTTCTGGAAT	2467
Db	2441	CGTTTGAAGAGCTCTCTTTTGTCTTGTGATATCAAGTTTCCCAAGAGATGGAATTTCTGGAAT	2500
Qy	2468	TTAAGTCTGTTTTCACAGAGACCTGGCGCCAGGAAGCTGTGTACCCAGGGGAAAG	2527
Db	2501	TCAGTCTGTTTTCACAGAGACCTGGCGCCAGGAAGTGTGTGTACCCAGGGGAAAG	2550
Qy	2528	TGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGAATTTCAACTATGTTG	2587
Db	2561	TGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGAATTTCAACTATGTTG	2620
Qy	2588	TCAGGGGCAATGCGCGTCTGCTGTAAATGGATGGCCCCGAAAGCCTGTTTGAAGGCA	2647
Db	2621	TCAGGGGCAATGCGCGTCTGCTGTAAATGGATGGCCCCGAAAGCCTGTTTGAAGGCA	2680
Qy	2648	TCTACACATTAAGAGTGTCTGTGTCATATGGAATATTAAGTGGGAAATCTTCTCAC	2707
Db	2681	TCTACACATTAAGAGTGTCTGTGTCATATGGAATATTAAGTGGGAAATCTTCTCAC	2740
Qy	2708	TTGTTGTAATTCCTTACCTGGCATTCCTGGTGTGATGCTAACTTCTACAACTGATTCAAA	2767
Db	2741	TGGTGTGAACCTTACCTGGCATTCCTGGTGTGATGCTAACTTCTACAACTGATTCAAA	2800
Qy	2768	ATGGATTTAAATGGATCAGCCATTTTATGTACAGAGAAATATACATTAATGCAAT	2827
Db	2801	GTGGATTTAAATGGATCAGCCATTTTATGTACAGAGAAATATACATTAATGCAAT	2860
Qy	2828	CCTCTGGCTTTTGAATCAAGGAAAGCCCATCTCTCCCTAACTTGTGCTGCTTTTGTAG	2887
Db	2861	CCTCTGGCTTTTGAATCAAGGAAAGCCCATCTCTCCCTAACTTGTGCTGCTTTTGTAG	2920
Qy	2888	GATGTAGCTGGCAGATGCAGAAAGCGATGTATCAGAAATGTGGATGGCGCTGTTTCGG	2947
Db	2921	GATGTAGCTGGCAGATGCAGAAAGCGATGTATCAGAAATGTGGATGGCGCTGTTTCGG	2980
Qy	2948	AATGCTCTCACCTTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGATTTGGGGCTAC	3007
Db	2981	AACATCTCATCTACCAAAACAGGCGCGCTTCAGCAGAGAGAGCGGGCTCAGAGCCGC	3040
Qy	3008	TCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACTTTAGTTTAAAGGACTTCAT	3067
Db	3041	CATGCCACAGGCGCCAGGT--GAAGATTCACAGAGAAAGATTTAGCGAGGAGCCCTTGA	3099
Qy	3068	CCCTTCCACCTATCCCTAACAGGCTGTAGATTTACCAAAACAGATTAATTTTCATCACTAAA	3127
Db	3100	CCCGCCCA-----CCCTAGCAGGCTGTAGAGCCGAGAGCCAGATTTAGCCTCGCCCT--G	3153
Qy	3128	AGAAATCTATTAATCAACTGCTGTTCACAGACTTTTCTCTAGAGCCGCTTCTGCTTTTA	3187
Db	3154	AGGAAGCCCTACAGGCGGTTGCTTCTGCTGAGCTTTTCTCTAGATGCTGTCTGCTTGA	3213
Qy	3188	CTCTTGTCTTCAAGGAGACTTTTGTAAATCAATCATCTCTGTACAGAGAGGAGGAGC	3247
Db	3214	CTC-----CAAAGTACTTCTATAAATCAAACTCTCTCTCGCACAGGGGAGAGCC	3266
Qy	3248	TGATAATGAATTTTGGAGCATTTGATCTGCATCCAGGCGCTTCTCAGCGCGCTTCTGAG	3307
Db	3267	ATAATGAGACTTGTGTTGGTGGCCCGCTACCCCTGGGGGGCTTTTCCAGGCCCCCAGGC	3326

Qy 3308 TGAATG-----TGTACCTGAAGTACAGTATATATCTTGTAAATACATAAAACAAA-- 3357  
Db 3327 TTGAGGGGAAAGCCATCTCTGAATATATAGTATATCTTGTAAATACGTTGAACAAACC 3386  
Qy 3358 -----ACATTTTCTAAGAGAGCTAAATGATTTTTT--AGTCTATGTTTTAAAT 3410  
Db 3387 AAACCCCGTTTTTGTCTAAGGAAAGCTAAATATGATTTTTTAAATAATCTATGTTTTAAAT 3446  
Qy 3411 AATATGTAAATTTTTTACGTATTTTAGTGATATATTTATGGTGGGAATAAAATTTCTAC 3470  
Db 3447 ACTATGTACTTTTTTCTATCTATTTAGTGATATATTTATGGATGGAATAAATTTCTAC 3506  
Qy 3471 TACAGAAAAA 3485  
Db 3507 TGTAGAAAAA 3521

RESULT 12  
AR005211  
LOCUS AR005211 3453 bp DNA linear PAT 04-DEC-1998  
DEFINITION Sequence 1 from patent US 5747651.  
ACCESSION AR005211  
VERSION AR005211.1 GI:3966090  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3453)  
AUTHORS Lemischka, I.R.  
TITLE Antibodies against tyrosine kinase receptor flk-1  
JOURNAL Patent: US 5747651-A 1 05-May-1998;  
FEATURES  
source Location/Qualifiers  
1. 3453  
/organism="unknown"

BASE COUNT 947 a 822 c 875 g 809 t  
ORIGIN  
Query Match 63.6%; Score 2227.2; DB 6; Length 3453;  
Best Local Similarity 79.6%; Pred. No. 0;  
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

Qy 35 GGGAGCCCGGGCTCCGGAGGCCATGCGCGGTGGC---GCGGACCGGGCACCCTGC 91  
Db 8 GGCTACCGCGCTCCGGAGGCCATGCGCGGTGGCGCAGCGAGCGCGGCTGC 67  
Qy 92 CGCTGCTGCTGTTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCCTG 151  
Db 68 TGCTGCTGTTGTTTGTGCAATGATTTGTGAGACCGTTACAAACCAAGACCTGCTG 127  
Qy 152 TGATCAAGTGTGTTTTTAATCAATATAAGAAATGATTCTAGTGGGAGTCAATCAT 211  
Db 128 TGATCAAGTGTGTTTTTAATCAATGATGATGAGAACATGGCTCATCAGCGGGAAGCCATCAT 187  
Qy 212 CATATCCCATGGTATCAGAAATCCCGGAGACCTCGGGTGTGCGTTGAGACCCAGAGCT 271  
Db 188 CGTACCGAATGGTGGAGGATCCCGAGAGACCTCCAGTGTACCCGAGCGGCCAGAGTG 247  
Qy 272 CAGGACAGTGTACGAAGCTCGCGTGTGGAAGTGGATGTATCTGCTTCCATCACACTGC 331  
Db 248 AAGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCGGAGTCTGGGTCCATCACCCCTGC 307  
Qy 332 AAGTGTGCTGATGCGCCCGGAGCAATTTCTGCTGCTGGGTCTTTAAGCACAGCTCC 391  
Db 308 AAGTGCAGTCCCGCCCGGAGGACCTTTCTGCTGCTGGGTCTTTAAGCACAGCTCC 367  
Qy 392 TGAATTGCCAGCCATTTTGAATTTACAAAACAGAGGAGTGTGTTTCCATGGTCAATTTGA 451  
Db 368 TGGGCTGCCAGCCGACCTTGTGATTTACAAACAGAGGATCGTTTCCATGGCCATCTGA 427  
Qy 452 AAATGACAGAAACCCAGCTGGAGATACCTACTTTTTTATTCAGAGTGAAGTACCAAT 511  
Db 428 ACGTGACAGAGACCCAGGAGGAGATACCTACTCCATATTTCAGAGCGAAGCGCAACT 487

Qy 512 ACACAATATTTTACAGTGTAGTATAAGAAATACCTCTGCTTTACACATTAAGAAGACCTT 571  
Db 488 ACACAGTACTTTTACAGTGTAGTATAAGAAATACCTCTGCTTTACACATTAAGAAGACCTT 547  
Qy 572 ACTTTAGAAAATGAAACACAGGAGCCCTGCTGCTATCTGATATCTGAGAGCGTTCCAGAGC 631  
Db 548 ACTTTAGAAAATGAAACACAGGAGTGCATCTGCTATCTCCAGAGGTGTTCCGAGC 607  
Qy 632 CGATCTGGAATGGTGTCTTTCCGATTTACAGGGGAAAGCTGTAAAGAAAGAAAGTCCAG 691  
Db 608 CCACCTGTGGAGTGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 667  
Qy 692 CTGTTGTTAAAGAGAGGAAAGTGTCTGATTAATTTTGGGACGGACATTAAGTGTCT 751  
Db 668 CTGTTGTCAGAAAGGAGGAAAGTGTCTGATGTTGTTCCGAAACAGACATCAGATGCT 727  
Qy 752 GTGCCAGAAATGAACCTGGGAGGAAATGCACAGGCTGTTCACATAGATCTAAATCAAA 811  
Db 728 GTGCTAGAAATGCACTGGGCGCGAATGCACCAAGCTGTTCAACATAGATCTAAACAGG 787  
Qy 812 CTCTCTCAGACACATTTGCCACAATTTATTTCTTAAAGTAGGGAAACCTTTATGGATAAGT 871  
Db 788 CTCTCTCAGACACACTGCCCCAGTTATCTCTGAAAGTGGGAAACCTTTGTGATCAGGT 847  
Qy 872 GCAAAGCTGTTCAATGTGAACCATGGAATTCGGGCTCACCTGGGAAATTTAGAAACAAAGCAC 931  
Db 848 GTAGGCCATCCATGTGAACCATGGAATTCGGGCTCACCTGGGAGCTGGAAGACAAAGCCC 907  
Qy 932 TCGAGGAGGCAACTACTTTGAGATGAGTACCTATTCAACAACACAGACATATCATACGGA 991  
Db 908 TGGAGGAGGCGAGCTACTTTGAGATGAGTACCTACTCCACAACAGGACCATGATTCGGA 967  
Qy 992 TTCTGTTGCTTTTGTATCATCAGTGGCAAGAAAGACACACCGGATACTACACTTGTTCCT 1051  
Db 968 TTCTGTTGCTTTTGTATCTGCTTCCGTGGGAGGAGACACACCGGATATTACACTGCTCT 1027  
Qy 1052 TTCTCAAGCATCCAGTCAATCAGCTTTGGTTACCATCTGAGAAAGGATTTATTAATG 1111  
Db 1028 COTCAAGCACCACCGGAGTGGTGGTACCATCTAGAAAAGGTTTATTAACG 1087  
Qy 1112 CTACCAATTCAGTCAAGATTAATGAATGACCAATATGAAGATTTGTTTCTCTGTC 1171  
Db 1088 CTACCAAGTCCGAAAGAGTATGAATGACCCGCTACGAAAGTCTGCTTCTCTAGTCA 1147  
Qy 1172 GGTTTAAAGCCTTACCACAAATCAGATGATGAGTGGACCTTCTCTCGAAATATTTCCCT 1231  
Db 1148 GGTTTAAAGCCTTACCACAAATCAGATGATGAGTGGATCTCTCTCAAGCCTCATTTCCCT 1207  
Qy 1232 GTGAGCAAAAGGCTTTGATAACGGATACAGCATATCCAAAGTGTTCGATCATAGCACC 1291  
Db 1208 GTGAACAGAGAGGCTTGGAGGATGGTACAGCATATCTAAATTTTGGATCATAGAACA 1267  
Qy 1292 AGCCAGGAGATATATATTTCCATGCAAAATGATGCTCCCAATTTACCAAAATGTTCA 1351  
Db 1268 AGCCAGGAGATATATATTTCTATGCAAAATGATGACCCGCTTACCAAAATGTTCA 1327  
Qy 1352 CGCTGAATATAAGAGGAAACCTCAAGTGTCTCGCAAGCATCGCAAGTTCAGGCGCTCT 1411  
Db 1328 CGCTGAATATAAGAGAAACCTCAAGTGTGATGATGCTCAAGTGTGCTCAGCAGCCGCTCT 1387  
Qy 1412 GTTCTCTCGGATGATACCCATTTACCATTTGGACCTGGAGAGTGTTCAGACAAGTCTC 1471  
Db 1388 GTTCTCTGATGCTACCCGCTTACCCCTTTGGACCTGGAGAGTGTTCGACAAATCTC 1447  
Qy 1472 CCACTGCAACAGAGATCAGAGAGGATCTGGAATAGAAAGGCTTAACAGAAAAGTGT 1531  
Db 1448 CCAATTTGACCGAGGAAATCCCAAGAGAGTGTGGAATAAAAAGGCTTAACAGAAAAGTGT 1507  
Qy 1532 TTGGACAGTGGTGTGAGCAGTACTCTAAACATGATGAAGCCATAAAAGGTTCTCTCG 1591  
Db 1508 TTGGCAGTGGTGTGAGCAGTACTCTAAATATGATGAGCCGCGGAAAGGCTCTCTCG 1567  
Qy 1592 TCAAGTGTGTCATACAAATTCCTTGGCACATCTTGTGAGAGCATCTTTTAACTCTC 1651



1568 TCAGATGTCGCGTACAAATCTATGGCGACGCTTGGCAAAACCATCTTTTAAACTCAC 1627  
1562 CAGGCCCTCCCTTTCATCAAGACACATCTCATCTATGTCACACAAATGGTGTGTC 1711  
1628 CAGGCCCTCCCTTTCATCAAGACACATCTCTCTATGTCGACCAATGGGCTCTGTC 1687  
1712 TCCTCTTCATTTGCTGTTTAAACCTGCTAAATTTGTCACAAGTACAAAAAGCAATTTAGGT 1771  
1688 TCCCTTCATTTGCTCTCATTTGTTGATCTGCCACAATATACAAAAGCAATTTAGGT 1747  
1772 ATGAAGCCAGCTACAGATGATACAGGTGACCGGCTCTCTCAAGTAATGAGTACTTCTACG 1831  
1748 ACAGAGTCAAGTCAAGATGATCCAGGTGACTGGCCCCCTGGATAACGAGTACTTCTACG 1807  
1832 TTGATTTACAGAGATATGAATATGATCTCAATGGGAGTTTCCAGAGAAAATTTAGAGT 1891  
1808 TTGATTTACAGGAGTATGAATATGATCTCAATGGGAGTTTCCGAGAGAGAACTTAGAGT 1867  
1892 TTGGGAAGGTACTAGGATCAGGTGCTTTTGGAAAAGTGTGAAGCAACAGCTTATGAA 1951  
1868 TTGGGAAGGTCTGGGCTCGCGCTTTCGGGAGGGTGTGAAGCCACCGCTTATGGCA 1927  
1952 TTAGCAAAACAGGAGTCTCAATCCAGGTTCGCGTCAAAATGCTGAAAAGAAAAGCAGACA 2011  
1928 TTAGTAAACGGGAGTCTCAATTCAGGTGCGGTGAAGATGCTAAAAGAGAAGCTGACA 1987  
2012 GCTCTGAAGAGAGGCACTCATGTCAGAACTCAAGATGATGACCCAGTGGGAGCCACG 2071  
1988 GCTGTGAAAAGAGCTCTCATGTGCGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047  
2072 AGAATATTGTGAACCTGCTGGGGCGTGCACACTGTCAGGACCAATTTACTTGTATTTTG 2131  
2048 ACAACATCGTAATCTGCTGGGGCGATGCACACTGTCAGGGCCAGTGTACTTGTATTTTG 2107  
2132 AATACTGTGCTATGCTGATCTCTCAACTATCTAAGAAGTAAAGAGAAAATTTTACA 2191  
2108 AATATTGTGCTATGCTGACCTCTCAACTCACTTAAGAAGTAAAGAGAGAATTTTACA 2167  
2192 GGACTTGGACAGAGATTTTCAAGAACACAAATTTTCACTTTTACCCCACTTTTCAATCAC 2251  
2168 GGACATGGACAGAGATTTTAAAGAAATATAATTTTCACTTTTACCCCACTTTTCCAGGCAC 2227  
2252 ATCAAAATCCAGATGCTGCTGTTCAAGAGAGTTCAGATACACCCGACCTCGATCAAA 2311  
2228 ATTCAATTCAGCATGCTGCTGTTTACAGAGAGTTCAGTTACACCCGCTTGTGATCAGC 2287  
2312 TCTCAGGCTTCATGGGAATTCATTTCACTCTGAAGATGAATTAATGAATAACCAAA 2371  
2288 TCTCAGGCTTCATGGGAATTCATTTCACTCTGAAGATGAATTAATGAATAACCAAA 2347  
2372 AAAGGCTG-----GAAGAGAGAGGAGTGAATTTGCTTACATTTTGAAGATCTTCTTT 2425  
2348 AGAGGCTGGCAGAGAAGAGAGAGAGAGATTTTGAAGTGTGCTGACCTTTTGAAGACCTCTT 2407  
2426 GCTTTGTCATATCAAGTTGCAAGAGATGGAATTTCTGGAATTTTAACTGCTGTTTACA 2485  
2408 GCTTTGCTACCAAGTGGCCAAAGGATGGAATTTCTGGAAGTTCAGTGTGTTGCCACA 2467  
2486 GAGACCTGGCCGAGGAAGTGTGTTGTCACCCACGGGAAGTGTGAAGATATGTGACT 2545  
2468 GAGACCTGGCAGCAGGAATGTTGTTGTCACCCACGGGAGGTGTGAAGATCTGTGACT 2527  
2546 TTGGATTGGCTCGAGATATCATGATGATTCACACTATGTTGTCAGGGCAATGCCGCTC 2605  
2528 TTGGACTGGCCGAGACATCTGAGCGACTCCAGCTACGTCGTGAGGGCAACCGCGC 2587  
2606 TGCTGTAAATGGATGGCCCCGAAAGCCTGTTTGAAGGATCTACACCAATTAAGAGTG 2665  
2588 TGCCGCTGAAGTGGATGGCCACCGAGAGCTTATTTGAAGGATCTACACCAATCAAGAGTG 2647  
2666 ATGCTGCTCATATGGAATATTACTGTGGGAATCTTCTCACTTGTGTTGTAATCCTTACC 2725

2648 ACCTCTGGTCTAGGCACTCTTCTCTGGAGATATTTTCACTGGGTGTGAACCTTACC 2707  
2726 CTGGCATTCGGTGTGATGCTAATCTTACAAACTGATTCAAAATGATTTAAAATGGATC 2785  
2708 CTGGCATTCCTCTCGACGCTAACTTCTATAAATGATTCAGAGTGGATTTAAAATGGAGC 2767  
2786 AGCCATTTTATCTCAGAGAAATATACATTAATGAATCAATCTCTGGCTTTTGCAT 2845  
2768 AGCCATTTCTATCCACAGAGAGATATATCTTTGTAATGAATCAATCTCTGGCTTTTGCAT 2827  
2846 CAAGAAAGCGGCATCTCCCTTAATTTGACTTCGTTTATAGGATGTCAGCTGGCAGATG 2905  
2828 CAAGAAAGCGGCATCTTCCCAACCTGACTTCATTTTATAGGATGTCAGCTGGCAGAGG 2887  
2906 CAGAAGAGCGATGATATCAGAATATGAGTGGCGTGTTCGGAAATCTCCTCACACCTACC 2965  
2888 CAGAAGAAG-----CATGTATCAGAACAATCCATCTCATCTACC 2924  
2966 AAAACAGCGGACCTTTCAGCAGAGAGATGGATTTGGGGCTACTCTCTCCGACGGCTCAGG 3025  
2925 AAAACAGCGGCGCTCTCAGCAGAGAGCGGGCTCAGAGC-CCAGTCGCCACACGCGCAGG 2983  
3026 TCGAAGATTCTGAGAGGAACAATTTAGTTTAAAGACTTTCATCCCTCCACCTATCCCTAA 3085  
2984 T-GAAGATTCACAGAGAAGAAGTTAGCGAGAGGGCTTGGACCCCGC-----CACCTAG 3038  
3086 CAGGCTGTAGATTACCAAAACAAAGATTAATTTTATCATCTATAAGAAAATCTATTATCAAC 3145  
3039 CAGGCTGTAGACCGCAGAGCAAGATTAGCTCTGCGCTCT--GAGGAAGCGCCCTACACGC 3096  
3146 TGCTCTTTCACAGAGACTTTTCTCTAGAGCGCTGCTGCTTTTCTTCTTCAAGAGGA 3205  
3097 CGTTGCTTCTGCTGGACTTTTCTCTAGATGCTCTGCTCCATTACTC-----CAAAAGTGA 3149  
3206 CTTTGTAAATCAATCATCTCTGCACAAAGCAGGAGGCTGATGAATGAACCTTTATTG 3265  
3150 CTTCTATAAATCAACCTCTCTCGCACAGCGGGAGAGCAATAATGAGACTTCTTGG 3209  
3266 GAGCATGATCTGCATCCAAAGCCTTCTCAGCGCGCTTTGAGTGAATTTGTACCTGAAG 3325  
3210 TGAGCGCCCTACCTCGGGGCTTTCCACGAGCTTGAGGGGAAAGCCATGTATCTGAAA 3269  
3326 TACAGTATATCTCTTAATACATATAACAAA-----AGCATTTTGTCTAGGAGAACG 3378  
3270 TATAGTATATCTCTGTAATACGTGAAACAAACCCGCTTTTGTCTAGGGAAGC 3329  
3379 TAATATGATTTT---AAGTCTATCTTTTAAATAATATGTAATTTTTCAGCTATTTAG 3436  
3330 TAATATGATTTTAAATAATCTATGTTTAAATACTATGTAATTTTTCATCTATTAG 3389  
3437 TGATATATTTTATGGTGGGAATAAATTTCTTACTACAGAAAATAAATAAATAAATAA 3496  
3390 TGATATATTTTATGGATGAAATAAATTTCTTACTGTAATAAATAAATAAATAAATAA 3449  
3497 AAAA 3500  
3450 AAAA 3453

## RESULT 13

AR071703  
LOCUS AR071703 3453 bp DNA linear PAT 18-FEB-2000  
DEFINITION Sequence 1 from patent US 5912133.  
ACCESSION AR071703  
VERSION AR071703.1 GI:7222591  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 3453)  
AUTHORS Lemischka, I. R.  
TITLE Method for isolating stem cells expressing flk-1 receptors  
JOURNAL Patent: US 5912133-A 1 15-JUN-1999;



FEATURES		Location/Qualifiers		Query Match		DB 6; Length 3453;	
source		1..3453		63.6%; Score 2227.2;		DB 6; Length 3453;	
BASE COUNT		947 a 822 c 875 g 809 t		Best Local Similarity 79.6%; Pred. No. 0;		Mismatches 2775; Conservative 0;	
ORIGIN							
Qy	35	GGGACCCCGGCTCCGGAGGCCATCGCGGCTGGC---	CGCGAGCGGGCACCGCTGC	91	11	11	11
Db	8	GGCTACCGCGCGCTCCGGAGGCCATCGCGGCTGGC	CGCGAGCGGGCACCGCGGCTGC	67	11	11	11
Qy	92	CGCTGCTGCTGTTTTCGCAATGATATTGGGACTATT	TACAAATCAAGATCGCTG	151	11	11	11
Db	68	TGCTGCTGTTGTTTGTGCTAGTATGATCTTGAGAC	CGTTACAAACCAAGACCTGCTG	127	11	11	11
Qy	152	TGATCAAGTGTGTTTAAATCAATCATAGAACAATG	ATCATAGTGGGGAAGTCATCAT	211	11	11	11
Db	128	TGATCAAGTGTGTTTAAATCAATCATAGAACAATG	ATCATAGTGGGGAAGTCATCAT	187	11	11	11
Qy	212	CATATCCATCGTATCAGAAATCCCGGAAGACCTCG	GGTGTGCTTGGAGCCGCCAGAGCT	271	11	11	11
Db	188	CGTACCGAATGGTCGAGGATCCCGAGAAACCTCC	AGTGTACCCGAGGCGCCAGAGTG	247	11	11	11
Qy	272	CAGGACAGTGTACGAAGCTCGGCTGTGGAAGTGA	TATCTGCTTCCCATCACACTGC	331	11	11	11
Db	248	AAGGACCGGTATATGAAGCGCCACCGTGGAGTGG	CGGAGTCTGGGTCCATCACCTGC	307	11	11	11
Qy	332	AAGTGTGCTGATGCTCCCGAGGAACATTTCTGCT	CTGCTGCTTTAAGCACAGCTCCC	367	11	11	11
Db	308	AAGTGTGCTGATGCTCCCGAGGAACATTTCTGCT	CTGCTGCTTTAAGCACAGCTCCC	367	11	11	11
Qy	392	TGAATGCCAGCCACATTTGATTTACAAACAGAGG	AGTGTCTTCCATGTCATTTGA	451	11	11	11
Db	368	TGGGTGCGCGCGACATTTGATTTACAAACAGAG	GAATGCTTCCATGTCATTTGA	427	11	11	11
Qy	452	AAATGACAGAAACCAAGCTGGAGAAATACCTACT	TTTATTTCAGAGTGAAGTACCAATT	511	11	11	11
Db	428	ACGTGACAGAGACCCGAGGAGAGAAATACCTACT	CCATATTCAGAGCGAAGCGCCACT	487	11	11	11
Qy	512	ACACAATATTTGTTTACGTAGTATAGAATAATAC	CCCTGCTTTACATATTAAGAAGCTTT	571	11	11	11
Db	488	ACACAGTACTGTTACAGTGAATTAAGAGATACAC	AGCTGATGTGCTTAAGGAGACCTT	547	11	11	11
Qy	572	ACTTTAGAAAATGGAACACAGGAGCGCCCTGGT	CTGCTATCTGAGACGCTTCCAGAGC	631	11	11	11
Db	548	ACTTTAGGAAGATGGAACACAGGAGCTGCTGCT	GTGCTTCCGAGGCTTCCGGAGC	607	11	11	11
Qy	632	CGATCGTGAATGGGTGCTTTGCGATTACAGGGG	AAAGCTGTAAGAAGAAAGTCCAG	691	11	11	11
Db	608	CCACTGTGGAGTGGGTGCTTCCGAGTCCACAGG	GAAGCTTAAGAAGAGGCGCTG	667	11	11	11
Qy	692	CTGTTCTTAAAAGGAGGAAAAGTCTTCATGAAT	TATTTGGGACGGACATTAAGTGTCT	751	11	11	11
Db	668	CTGTTCTGAGAGGAGGAAAAGTCTTCATGATGT	TTCGGAACAGACATCAGATGCT	727	11	11	11
Qy	752	GTGCCAGAAATGAATGGGAGGGAATGCACAGGCT	GTTCACAATAGATCTTAATCAAA	811	11	11	11
Db	728	GTGCTAGAAATGCATGGGCGCGGAATGCACCAAG	CTGTTCACCATAGATCTTAACCCAGG	787	11	11	11
Qy	812	CTCCTCAGACCACATTTGCCACAAATTTTCTTAA	AGTATGGGGAACCTTATGATAGGT	871	11	11	11
Db	788	CTCCTCAGACACACTGCCCAATTTATCTGAAAGT	GGGGAAACCTCTGTGATCAGGT	847	11	11	11
Qy	872	GCAAGCTGTTCATGTCAACCATGGATTCCGGGCT	CACCTGGGGAATTAGAAAACCAAGCAC	931	11	11	11
Db	848	GTAAAGCCATCCATGTGAACCATGGATTCCGGCT	CACCTGGGAGCTGGAAGACAAAGCCC	907	11	11	11
Qy	932	TCGAGGAGGCGCACTACTTTGAGATGAGTACCTAT	TTCACAAACACAGAACTATGATACGGA	991	11	11	11
Db	908	TGAGGAGGCGCACTACTTTGAGATGAGTACCTACT	CCACAAACAGGACCATGATTCGGA	967	11	11	11

Qy	992	TTCTGTTTGTCTTTTGTATCATCAGTGGCAAGAAAC	GACACCGGATACATACACTTGTCTCT	1051	11	11	11
Db	968	TTCTTGTGGCTTTTGTCTTCCGTGGAGAGACGACAC	CGGATATTACACCTGCTCTT	1027	11	11	11
Qy	1052	CTTCAAGACATCCAGTCAATCAGCTTTGGTTTACC	ATCGTAGGAAAGGATTTATAATG	1111	11	11	11
Db	1028	CTTCAAGACATCCAGTCAATCAGCTTTGGTTTACC	ATCGTAGGAAAGGATTTATAATG	1087	11	11	11
Qy	1112	CTACCAATTCAGTGAAGATTATGAATTTGACCAAT	ATATGAAGAGTTTCTTTTCTGTCA	1171	11	11	11
Db	1088	CTACCAATTCAGTGAAGATTATGAATTTGACCAAT	ATATGAAGAGTTTCTTTTCTGTCA	1147	11	11	11
Qy	1172	GGTTTAAAGCCTACCCACAAATCAGATCTAGCTGC	ACCTTCTCGAAATCATTTCCCTT	1231	11	11	11
Db	1148	GGTTTAAAGCCTACCCACAAATCAGATCTAGCTGC	ACCTTCTCGAAATCATTTCCCTT	1207	11	11	11
Qy	1232	GTGAGCAAAAGGCTTGTGATAACGGATACAGCAT	ATATCCAAAGTTTGGCAATCAAGACCC	1291	11	11	11
Db	1208	GTGAGCAAAAGGCTTGTGATAACGGATACAGCAT	ATATCCAAAGTTTGGCAATCAAGACCC	1267	11	11	11
Qy	1292	AGCCAGGAGATATATATTTCCATGCGAGAAATGA	TGATGCCCAATTTACCAAAATCTCA	1351	11	11	11
Db	1268	AGCCAGGAGATATATATTTCCATGCGAGAAATGA	TGATGCCCAATTTACCAAAATCTCA	1327	11	11	11
Qy	1352	CGCTCAATATAGAAGGAACCTCAAGTCTCGAGAAC	CATCGGCAAGTTCAGGCGTCTCT	1411	11	11	11
Db	1328	CGCTCAATATAGAAGGAACCTCAAGTCTCGAGAAC	CATCGGCAAGTTCAGGCGTCTCT	1387	11	11	11
Qy	1412	GTTCCTCGGATGGATACCCATTTACCATTCTTGG	ACCTTGGAGAGTGTTCAGACAAGTCTC	1471	11	11	11
Db	1388	GTTCCTCGGATGGATACCCATTTACCATTCTTGG	ACCTTGGAGAGTGTTCAGACAAGTCTC	1447	11	11	11
Qy	1472	CCAATGTCACAGAGAGATCAGAGAGTCTGGAAT	AGAAGGCTTAACAGAAAAGTGT	1531	11	11	11
Db	1448	CCAATGTCACAGAGAGATCAGAGAGTCTGGAAT	AGAAGGCTTAACAGAAAAGTGT	1507	11	11	11
Qy	1532	TTGGACAGTGGTGTGCGACGACTTAAACATGAGT	GAAGCATAAAGGCTTAACAGAAAAGTGT	1591	11	11	11
Db	1508	TTGGACAGTGGTGTGCGACGACTTAAACATGAGT	GAAGCATAAAGGCTTAACAGAAAAGTGT	1567	11	11	11
Qy	1592	TCAAGTGTGTGATACAAATTCCTTTGGCACATCT	TTGTGAGACGATCCTTTTAAACCTCTC	1651	11	11	11
Db	1568	TCAAGTGTGTGATACAAATTCCTTTGGCACATCT	TTGTGAGACGATCCTTTTAAACCTCTC	1627	11	11	11
Qy	1652	CAGGCCCCCTTCCCTTTTATCCCAAGACAACTCT	ATTCATGCAACAATTTGGTGTCTC	1711	11	11	11
Db	1628	CAGGCCCCCTTCCCTTTTATCCCAAGACAACTCT	ATTCATGCAACAATTTGGTGTCTC	1687	11	11	11
Qy	1712	TCTCTTTCATTTGCTGTTTTTAAACCTTCTAAT	TTGTACAAAGTACAAAAGCAATTTAGGT	1771	11	11	11
Db	1688	TCTCTTTCATTTGCTGTTTTTAAACCTTCTAAT	TTGTACAAAGTACAAAAGCAATTTAGGT	1747	11	11	11
Qy	1772	ATCAAAAGCCAGTACAGATGGTACAGGTGACCG	CTCCTCAGATAATGAGTACTTCTACG	1831	11	11	11
Db	1748	ATCAAAAGCCAGTACAGATGGTACAGGTGACCG	CTCCTCAGATAATGAGTACTTCTACG	1807	11	11	11
Qy	1832	TTGATTTACAGAAATATGAATATGATCTCAATTA	AGTGGAGTTTCCAAAGAGAAATTTAGGT	1891	11	11	11
Db	1808	TTGATTTACAGAAATATGAATATGATCTCAATTA	AGTGGAGTTTCCAAAGAGAAATTTAGGT	1867	11	11	11
Qy	1892	TTGGAAGGTACTAGGATCAGGTGCTTTTGGAAAG	GTGATGACCGCAACAGCTTATGGAA	1951	11	11	11
Db	1868	TTGGAAGGTACTAGGATCAGGTGCTTTTGGAAAG	GTGATGACCGCAACAGCTTATGGAA	1927	11	11	11
Qy	1952	TTAGCAAAACAGGAGTCTCAATCCAGGTTCGCGT	CAAAATGCTGAAAGAAAAGCAGACA	2011	11	11	11
Db	1928	TTAGCAAAACAGGAGTCTCAATCCAGGTTCGCGT	CAAAATGCTGAAAGAAAAGCAGACA	1987	11	11	11
Qy	2012	GCCTCTGAAAGAGGCGCACTCATGTCAGAACTCA	AGATGATGACCCAGCTGGGAAGCCACG	2071	11	11	11
Db	1988	GCCTCTGAAAGAGGCGCACTCATGTCAGAACTCA	AGATGATGACCCAGCTGGGAAGCCACG	2047	11	11	11

QY 2072 AGAATATTGTGAACCTGCTGGGGCGTGCACACTGTCAAGACCAATTTACTGATTTTG 2131  
Db 2048 ACAACATCGTAATCTGCTGGGGCATGCACACTGTCAAGGCCAGTGACTGATTTTG 2107  
QY 2132 AATAGTGTGCTATGCTGATCTTCAACTATCTAAGAAGTAAAGAGAAAATTTTACA 2191  
Db 2108 AATATGTTGCTATGCTGACCTTCCCTCACTAGTAAAGTAAAGAGAAATTTTACA 2167  
QY 2192 GGACTTGGACAGAGATTTTCAAGGAACACAATTTTACAGTTTACCCCACTTTCCATCAC 2251  
Db 2168 GGACATGGACAGAGATTTTAAAGAACATATTTTACCTTCTTACCTACTTTTCCAGGCAC 2227  
QY 2252 ATCCAAATTCAGCATGCTGCTTCAAGAGAAAGTTTCAAGATACACCGGACTCGGATCAAA 2311  
Db 2228 ATTCAAATTCAGCATGCTGCTTCAAGAGAAAGTTTCAAGTTACACCGCCCTTGGATCAGC 2287  
QY 2312 TCTCAGGCTTTCATGGGAATTCATTTCACTCTGAAGATGAATTAATTAAGAAACCAA 2371  
Db 2288 TCTCAGGCTTTCATGGGAATTCATTTCACTCTGAAGATGAATTAATTAAGAAACCAA 2347  
QY 2372 AAAGCTG-----GAAGAGAGAGGAGCTTGAATCTGCTTACATTTGAAGATCTTCTT 2425  
Db 2348 AGAGCTGGCAGAGAGAGAGAGATTTGAACGTGCTGAGCTTTGAAGACCTCTCTT 2407  
QY 2426 GCTTTGCATATCAAGTTGCCAAAGGAATTTCTGGAATTTAAGTCTGCTGTTTACA 2485  
Db 2408 GCTTTGCGTACCAAGTGGCCAAAGCATGGAATCTCTGAGTTCAAGTCTGCTGCCACA 2467  
QY 2486 GAGACTGGCCCGCAGGACGTGCTTGCACCCAGGGAAGTGGTGAAGATATGACT 2545  
Db 2468 GAGACTGGCCCGCAGGAGATGTTGGTCAACCAGGGAAGTGGTGAAGATATGACT 2527  
QY 2546 TTGGAATGCTGAGATATCATGACTGATTTCAACTATGTTGTCAGGGCAATGCCGTC 2605  
Db 2528 TTGGAATGCTGAGACATCTCTGAGCGACTCAGCTAGCTGTCAGGGCAACGACGCG 2587  
QY 2606 TGCTGTAAATGGATGGCCCGCAGGAGCTGTTTGAAGCATCTACACATTAAGAGTG 2665  
Db 2588 TGCCGTGAGTGGATGGCCCGCAGGAGCTTATTTGAAGGATCTACACATCAAGAGTG 2647  
QY 2666 ATGCTGTGCTATATGAATATTTACTGTGGGAATCTTCTCACTTGGTGAATCTTACC 2725  
Db 2648 ACGTCTGCTCTACGGCATCTCTCTGGAGATATTTTCACTGGTGTGAACCTTACC 2707  
QY 2726 CTGCAATCCGCTGATCTACTTCTACAACTGATTCAAATGGAATTAATGGATC 2785  
Db 2708 CTGCAATCCGCTGATCTACTTCTATAAACTGATTCAGAGTGGATTTAAATGGAGC 2767  
QY 2786 AGCCATTTTACTCAGAGAAATATATACATTAATGAATCACTCTGCTGGGCTTTTGACT 2845  
Db 2768 AGCCATTTTACTCAGAGAAATATATCTTTGTAATGAATCTCTGCTGGGCTTTTGACT 2827  
QY 2846 CAAGAAACGGCATCTTCCCTAATTTGACTTCTTTTAGATGTCTAGCTGGCAGATG 2905  
Db 2828 CAAGAAACGGCATCTTCCCTAATTTGACTTCTTTTAGATGTCTAGCTGGCAGATG 2887  
QY 2906 CAGAAAGACGATGATCAGATGGGATGGCCGCTGTTTCGAAATGCTCTACACCTACC 2965  
Db 2888 CAGAAAG-----CATGTATCAGAACTCCATCCATCTACC 2924  
QY 2966 AAAACAGCGGACCTTTTCAAGAGAGATGGATTTTGGGGCTACTCTCTCCGAGGCTCAGG 3025  
Db 2925 AAAACAGCGGACCTTTTCAAGAGAGGCGGCTCAGAGC-CAAGTCCGACAGCGGCGAG 2983  
QY 3026 TCGAAGATTCGTAGAGAACAAATTTAGTTTAAAGGACTTCACTCCCTCCACCTATCCCTAA 3085  
Db 2984 T-GAAGATTCAGAGAAAGAAAGTTAGCGAGGAGGCTTGGACCCGCG-CAACCTAG 3038  
QY 3086 CAGGCTGTAGATTACCAAAACAAGATTAATTTATCATCTATAAAGAAATCTATTATCAAC 3145  
Db 3039 CAGGCTGTAGACCGCAGAGCAAGATTTAGCCTCGCCTCT--GAGGAAGCGGCTTACAGCG 3096  
QY 3146 TGCTGCTTACCAGACTTTTCTCTAGAGCGGCTCTGCGCTTACTCTTGTTCCTTCAAGGGA 3205

Db 3097 CGTTGCTTCCGTGACTTTTCTCTAGATGCTGTCTGCCATTACTC-----CAAAGTGA 3149  
QY 3206 CTTTGTGTAATAAATCAATCATCTCTGTACAGGAGGAGGAGCTGATTAATGAACCTTATTG 3265  
Db 3150 CTTTGTGTAATAAATCAATCATCTCTGTACAGGAGGAGGAGCTGATTAATGAACCTTATTG 3209  
QY 3266 GAGCATGTACTGCATCCCAAGGCTTCTCAGGCGGCTTGTAGTCAATTTGTCTACTGAAG 3325  
Db 3210 TGAGCCCGCTACCTCGGGGCTTTTCCAGAGCTTGAAGGGAAGGCATGTATCTGAAA 3269  
QY 3326 TACAGTATATTTCTGTAATAATACATAAACA-----AGCATTTTGTCTAAGGAGAAGC 3378  
Db 3270 TATAGTATATTTCTGTAATAATACGTGAACAACAACCAACCCGTTTTTGTCTAAGGAGAAGC 3329  
QY 3379 TAATATGATTTTTT--AAGTCTATGTTTTTAAATAATAATGATAATTTTTCAGCTATTATG 3436  
Db 3330 TAAATATGATTTTTTAAATAATCTATGTTTTTAAATAATGATAATTTTTCATCTATTATG 3389  
QY 3437 TGATATATTTTATGGGTGGGAATAAATTTCTACTACAGAAAAAATAAATAAATAA 3496  
Db 3390 TGATATATTTTATGGATGGAATAAATACTTTCTACTGTAAAAAATAAATAAATAA 3449  
QY 3497 AAAA 3500  
Db 3450 AAAA 3453

RESULT 14  
125169  
LOCUS 125169 3453 bp DNA linear PAT 07-OCT-1996  
DEFINITION Sequence 1 from patent US 5548065.  
ACCESSION 125169  
VERSION 125169.1 GI:1605039  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 3453)  
AUTHORS Lemischka, I.R.  
TITLE Tyrosine kinase receptor human flk-2-specific antibodies  
JOURNAL Patent: US 5548065-A 1 20-AUG-1996;  
FEATURES Location/Qualifiers  
source 1..3453  
/organism="unknown"  
BASE COUNT 947 a 822 c 875 g 809 t  
ORIGIN

Query Match. 63.6%; Score 2227.2; DB 6; Length 3453;  
Best Local Similarity 79.6%; Pred. No. 0;  
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;

QY 35 GGGGACCCGGGCTCGGAGGCCATCGCGCGTTGGC--GCAGCAGCGGACCCGCTG 91  
Db 8 GGCTACCGCGCTCGGAGGCCATCGCGCGTTGGCGCAGCGCAGCGCGGCTG 67  
QY 92 CGCTGCTGCTGTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCGTG 151  
Db 68 TGCTGCTGTTGTTTGTGCAATGATTTCTGAGACCGTTTACAAACCAAGACCTG 127  
QY 152 TGATCAAGTGTGTTTAAATCAATCAATGAACATGATTCATCAGTGGGGAAGTCAT 211  
Db 128 TGATCAAGTGTGTTTAAATCAATGATGAGAACAATGGCTCATCGCGGGAAGGCCAT 187  
QY 212 CATATCCCATGTATCAGAAATCCCGGAAGACCTCGGCTGCTGCTTGTAGACCCAGAGCT 271  
Db 188 CGTACCGAATGTGCGAGGATCCCGAGAGACCTCCAGTGATCCCGGAGCGCCAGAGT 247  
QY 272 CAGGACAGCTGTACGAAGCTGCGCTGTGGAGTGTATCTGCTTCCATCAGACTGC 331  
Db 248 AAGGACGCTATATGAAGCGGCCACCGTGGAGTGGCCGAGTCTGGTCTCATCCCTGC 307  
QY 332 AAGTCTGTGCTGATGCCCCCAGGGAACATTTCTGCTCTGCGGTCTTTAAGCAGAGCTCC 391



Qy	2546	TTGGATTGGCTCGAGATATCATGAGTATCTCCAACTATGTTGTCAGGGGCAATGCCCGTC	2605
Db	2528	TTGGACTGGCCGGAGACATCTGAGCGACTCCAGCTACGTCGTGAGGGGCAACGCACGGC	2587
Qy	2606	TGCCGTGAAATGGATGGCCCGCCGAAAGCCCTGTTTGAAGGCATCTACACCATTAAGAGTG	2665
Db	2588	TGCCGGTGAAGTGGATGGCACCAGAGAGCTTAATTTGAAGGATCTACACAATCAAGAGTG	2647
Qy	2666	ATGCTCGTCATATGGAATATTACTGTGGGAATCTTCTCAGTCTGGTGCTGAATCCCTTACC	2725
Db	2648	ACGTCTGGTCTACGGCATCTCTCTGGGAGATATTTTTCACGTGGGTGGAACCCCTTACC	2707
Qy	2726	CTGSCATTCCGGTTGATGCTAACTCTTACAAACTGATTCAAAATGGATTAAATATGGATC	2785
Db	2708	CTGSCATTCTGTGCAGCCTAACTTCTATAAACTGATTCAGATGGGATTTTAAATGGAGC	2767
Qy	2786	AGCCATTTTATGCTACAGAGAAATATACATTTATATGCAATCCTGCTGGGCTTTTGGACT	2845
Db	2768	AGCCATTCTATGCCACAGAAGGATATACTTTGTAAATGCAATCCTGCTGGGCTTTTGGACT	2827
Qy	2846	CAAGGAAACGGCCATCTTCCCTAAATTGACTTCGTTTTTAAAGATGTCAGCTGCGCAGATG	2905
Db	2828	CAGGAGCGGCCATCTCTCCCACTGACTTCATTTTTTAAAGATGTCAGCTGGCAGAGG	2887
Qy	2906	CAGAAGAGCGATGTATCAGAAATGTGGATGGCCGTGTTTCGGAATGTCCTCACACCTACC	2965
Db	2888	CAGAAGAG-----CATGTATCAGAATCCATCCATCTACTACC	2924
Qy	2966	AAACAGCGCACCTTTTCAGCAGAGATGATTTGGGGCTACTCTCTCCCGCAGGCTCAGG	3025
Db	2925	AAACAGCGCGCCCTTCAGCAGAGCGGGCTCAGAGC--CCAATGCGCCACAGCGCCAGG	2983
Qy	3026	TCGAAGATTTCGTAGAGGAACAATTTAGTTTTTAAAGACTTCATCCCTCCACCTATCCCTAA	3085
Db	2984	T-GAAGATTACAGAGAAAGAAGTTACGAGGAGGCCCTTGGACCCCG-----CACCTAG	3038
Qy	3086	CAGGCTCTAGATTACCAAAACAAGATTAATTTTATCACTAAAAGAAATCTATTATCAAC	3145
Db	3039	CAGGCTGTAGCCGCAGAGCCAAGATTAGCCTCGCCTCT--GAGGAAGCCCTACAGCG	3096
Qy	3146	TGCTGCTTCCACAGACTTTTCTCTAGAACCGCTCTGCTTTTACTCTTTGTTTCAAAGGGA	3205
Db	3097	CGTTGCTTCGCTGAGCTTTTCTCTAGATGCTGTCTGCCATTACTC-----CAAGTGA	3149
Qy	3206	CTTTTGTAAATCAATCATCTGTTCACAGCGAGGAGCTGTAATGAATCTTTATTTG	3265
Db	3150	CTTCTATATAAATCAAACTCTCTCGCACAGGGGGAGAGCCAATAATGAGACTTGTGTG	3209
Qy	3266	GAGCATTGTATCTGCATCCAAAGCCCTTCTCAGGCGCGCTTCAGTGAATTTGTTACCTGAAG	3325
Db	3210	TGAGCCCGCTACCTCGGGGCGCTTCCAGAGCTTCAGGGGAAGCCATGTATCTGAAA	3269
Qy	3326	TACAGTATATCTGTGTAATACATAAAACAAA-----AGCATTTTGGCTAAGGAGAAGC	3378
Db	3270	TATAGTATATCTTGTAAATACGTGAACCAAAACCAAAACCCCGTTTTCCTAAGGGAAGC	3329
Qy	3379	TAAATGATTTTTT--AAGTCTATGTTTTTAAATAATATGCTAAATTTTTTTCAGCTATTTAG	3436
Db	3330	TAAATATGATTTTTTAAATAATCTATGTTTTTAAATACATATGAACTTTTTCATCTATTAG	3389
Qy	3437	TGATATATTTTATGGGTGGGAATAAAATTTCTACTACAGAAAAAATAAAAAAAAAAAAAA	3496
Db	3390	TGATATATTTTATGGATGGGAATAAACTTCTACTGTAAAAAATAAAAAAAAAAAAAA	3449
Qy	3497	AAAA 3500	
Db	3450	AAAA 3453	

RESULT	15
I40600	
LOCUS	
I40600	DNA
3453 bp	linear
PAT	13-MAY-1997

DEFINITION	Sequence 1 from patent US 5621090.									
ACCESSION	I40600									
VERSION	I40600.1 GI:2082892									
KEYWORDS	Unknown.									
SOURCE	Unknown.									
ORGANISM	Unknown.									
REFERENCE	Unclassified.									
AUTHORS	1 (bases 1 to 3453)									
TITLE	Lemischka,I.R.									
JOURNAL	Nucleic acids encoding soluble human FLK-2 extracellular domain									
FEATURES	Patent: US 5621090-A 1 15-APR-1997;									
source	Location/Qualifiers									
	1..3453									
BASE COUNT	947 a 822 c 875 g 809 t									
ORIGIN										
	Query Match 63.6%; Score 2227.2; DB 6; Length 3453;									
	Best Local Similarity 79.6%; Pred. No. 0;									
	Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;									
Qy	35	GGGACCCCGGGCTCCGAGGCCATCGCGCTTGGC---	GC	CGACGCGGGCACCGTGC	91					
Db	8	GGCTACCGCGCTCCGAGGCCATCGCGCTTGGCGAGCGACGCCGGCGGTGC	67							
Qy	92	CGCTGCTCGTGTGTTTTCTGCAATGATATTTGGGACTATTACAAATCAAGATCTGCCTG	151							
Db	68	TGCTGCTGTGTGTTTTCTGASTAATGATCTTGAGACCGTTACAAACCAAGACCTGCCTG	127							
Qy	152	TGATCAAGTGTGTTTTAATCAATCATAGAACAATGATTCATCAGTGGGGAAGTCATCAT	211							
Db	128	TGATCAAGTGTGTTTTAATCAGTCNTCAGAACATGGCTCATCAGCGGGAAAGCCATCAT	187							
Qy	212	CATATCCCATGGTATCAGATPCCCGGAAGACCTCGGGTGTGGCTTGAGACCCACAGAGCT	271							
Db	188	CGTACCAGATGGTTCGAGGATCCCGAAGACCTCCACGTACCCCGAGCGCCAGAGTG	247							
Qy	272	CAGGACAGTGTACGAAGCTCGCGCTCTGGAAGTGATGTATCTGCTTCATCACACTGC	331							
Db	248	AAGGACGGTATATGAAGCGGCCACCGTGGAGGTGGCGAGTCTGGGTCCATCACCTCGC	307							
Qy	332	AAGTGTGGTGCATGCCCGCAGGAACAATTCCTGCTCTGGGTCTTTAAGCAGAGCTCCC	391							
Db	308	AAGTGCAGCTCGCCACCCAGGGGACCTTCCTGCTCTGGTCTTTAAGCAGAGCTCCC	367							
Qy	392	TGAATTCGACCCACATTTTGATTTACAAACAGAGAGAGTTGTTTCATGGTCATTTTGA	451							
Db	368	TGGCTGCCAGCCGACCTTTGATTTACAAACAGAGGAATCGTTTCCATGGCATCTTGA	427							
Qy	452	AAATGCAGAAAACCAAGCTGGAGAAATACCTACTTTTATTCAGAGTGAAGCTACCAATT	511							
Db	428	ACGTGCACAGAGACCCAGCAGGAGAAATACCTACTCCATATTCAGAGCGAAGCGCCAACT	487							
Qy	512	ACACAATATTGTTTACAGTGAGTATAGAATAACCCCTGCTTTACATTAAGAAGACCTT	571							
Db	488	ACACAGTACTGTTCCACAGTGAATGTAAGATACACAGCTGTATGTCTAAGGAGACCTT	547							
Qy	572	ACTTTAGAAAAATGGAACACAGGACCCCTCGTCTGCATATCTGAGAGCGTTCCAGAGC	631							
Db	548	ACTTTAGGAAGATGGAAACACGAGATGCATCTCTGCTATCTCCGAGGGGTGTTCCGGAGC	607							
Qy	632	CGATCGTGAATGGGTGCTTTTGGCATTCAGGGGGAAGCTGTAAAGAGAAGAAGTCCAG	691							
Db	608	CCACTGTGGAGTGGGTGCTCTGCAGCTCCACAGGGGAAGAGCTGTAAGAAGAAGGCCCTG	667							
Qy	692	CTGTTGTTAAAAAGGAGGAAAAAGTGTCTCATGAATTTATTTGGGACCGGACATGAAGTGC	751							
Db	668	CTGTTGTGAGAAAGGAGGAAAAAGTACTTTCATGAGTTGTTCCGAACACAGATCAGATGCT	727							
Qy	752	GTGCCAGAAATGAAGTGGGACGGAATGCACACAGGCTGTTTACAAATAGATCTTAATCAA	811							
Db	728	GTGCTAGAAATGCATGGCGGGGAATGACCAAGCTGTTCCACCATATGATCTTAACACAGG	787							



```
Db 2925 AAAACAGCGGCCCTCAGCAGAGAGCGGGCTCAGAGC-CCAGTCGCCACAGCGCCAGG 2983
Qy 3026 TCGAAGATTCTAGAGGAACAATTAGTTTAAAGGACTTTCCTCCCTCCACCTATCCCTAA 3085
Db 2984 T-GAAGATTACACAGAGAAAGAGTTAGCGAGAGGCGCTTGACCCCGC----CACCCTAG 3038
Qy 3086 CAGGCTGTAGATTACCAAAACAAGATTAAATTTCTACTATAAAGAAAATCTATTATCAAC 3145
Db 3039 CAGGCTGTAGACCGCAGAGCGCAAGATTAGCTCGCCTCT--GAGGAAGCGCCCTACAGCG 3096
Qy 3146 TGCTGCTTCACACAGACTTTCTCTAGAGCGGCTCTGCGTTTACTCTGTTTTCAAAGGA 3205
Db 3097 CGTTGCTTCGCTGGACTTTCTCTAGATGCTGTGCTGCCATTACTC-----CAAAGTGA 3149
Qy 3206 CTTTGTAAATCAATCATCTCTCACAGGAGGAGGAGCTGATAATGAACCTTTATTG 3265
Db 3150 CTTCTATAAATCAAACTCTCCTCGCACAGCGGGAGAGCCAAATATGAGACTTGTGG 3209
Qy 3266 GAGCATTGATCTGCATCCAAGGCCTTCTCAGCGCGGCTTGAGTGAATGTGTACTGAAG 3325
Db 3210 TGAGCCCGCCTACCTCGGGGCCCTTCCACGAGCTTGAGGGGAAAGCCATGTATCTGAA 3269
Qy 3326 TACAGTATATCTTGTAAATACATAAAACAAA-----AGCATTTGCTAAGGAGAAGC 3378
Db 3270 TATAGTATATCTTGTAAATACGTGAACAAACCAACCCCGCTTTTGTCTAAGGAAAGC 3329
Qy 3379 TAATATGATTTTTT--AAGCTATGTTTAAATATATGTAATTTTTCAGCTATTAG 3436
Db 3330 TAATATGATTTTAAATAATCTATGTTTAAATACTATGTAATTTTTCATCTATTAG 3389
Qy 3437 TGATATATTTATGGGTGGGAATAAATTTCTACTACAGAAAAAATAAAAAAAAAAAAAA 3496
Db 3390 TGATATATTTATGGATGGAATAAATTTCTACTGTAAAAAATAAAAAAAAAAAAAA 3449
Qy 3497 AAAA 3500
Db 3450 AAAA 3453
```

Search completed: May 25, 2003, 09:19:07  
Job time : 8785.17 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 04:06:12 ; Search time 704.832 Seconds  
(without alignments)  
11185.997 Million cell updates/sec

Title: US-09-919-408-3

Perfect score: 3501

Sequence: 1 CGAGGCGCATCCGAGGCT.....AAAAAAAAAAAAAAAAAAAA 3501

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_101002.\*

1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3501	100.0	3501	16	AAQ81013
2	3501	100.0	3501	16	AAQ79069
3	3501	100.0	3501	17	AAT38734
4	3501	100.0	3501	18	AAT72117
5	3501	100.0	3501	20	AAQ77515
6	3499.4	100.0	3501	14	AAQ53503
7	3499.4	100.0	3501	14	AAQ40915
8	3497.8	99.9	3501	14	AAQ35250
9	3474.4	99.2	3476	16	AAQ91536

10	3470.2	99.1	3475	16	AAQ00802	Human Flk2/flt3 ty
11	3086.4	88.2	3120	14	AAQ49756	PTK gene LptK25.
12	3083.2	88.1	3120	16	AAQ03096	protein tyrosine-k
13	2875.8	82.1	2949	19	AAV39041	Human receptor typ
14	2868.4	81.9	2958	19	AAV39040	Human receptor typ
15	2868.4	81.9	2958	19	AAV39042	Human receptor typ
16	2846	81.3	2982	19	AAV39039	Human receptor typ
17	2835.4	81.0	2978	19	AAV39038	Human receptor typ
18	2273	64.9	3521	16	AAQ00801	Flk2/flt3 tyrosine
19	2227.2	63.6	3453	16	AAQ81012	Human foetal prot
20	2227.2	63.6	3453	17	AAT38733	Human foetal liver
21	2227.2	63.6	3453	18	AAT72118	Murine flk-2 recep
22	2227.2	63.6	3453	20	AAQ77514	Murine flk-2 CDNA.
23	2225.6	63.6	3453	14	AAQ53502	Murine flk-2 CDNA.
24	2225.6	63.6	3453	14	AAQ35249	Murine flk-2 CDNA.
25	2225.6	63.6	3453	16	AAQ79068	Mouse flk-2 CDNA.
26	2222.4	63.5	3453	13	AAQ29954	Murine flk-2 CDNA.
27	2222.4	63.5	3453	14	AAQ40914	Murine flk-2 CDNA.
28	1545.8	44.2	2247	19	AAV55294	Sequence pMON32390
29	625.4	17.9	1894	15	AAQ54036	Flk-2ws gene. Mus
30	330.4	9.4	332	13	AAQ29955	Eztracellular doma
31	324.2	9.3	3992	23	AAQ79666	DNA encoding novel
32	324.2	9.3	3992	24	ABK48105	Human macrophage c
33	322.6	9.2	3069	21	AAZ4718	Bovine c-Kit bk-1
34	317	9.1	5098	22	AAQ13425	Murine Kit/stem ce
35	313.2	8.9	5084	19	AAV20443	Human c-kit oncoge
36	313.2	8.9	5084	22	AAQ13426	Human Kit/stem cel
37	313.2	8.9	5084	24	ABL64113	Breast cancer rela
38	313.2	8.9	5084	24	ABL68085	Ovary cancer relat
39	308.4	8.8	2919	20	AAV80887	Hampshire porcine
40	308.2	8.8	6390	23	AAQ79665	DNA encoding novel
41	308.2	8.8	6390	23	AAQ84936	DNA encoding novel
42	302.6	8.6	4054	11	AAQ06869	Sequence encodes P
43	302.6	8.6	6378	24	ABL62372	Colon adenocarcino
44	302.6	8.6	6378	24	ABL68519	Kidney cancer rela
45	302.6	8.6	6378	24	ABK35520	Human endometrial

#### ALIGNMENTS

RESULT 1	
AAQ81013	
ID	AAQ81013 standard; cDNA; 3501 BP.
XX	AAQ81013;
AC	AAQ81013;
XX	
DT	18-AUG-1995 (first entry)
XX	
DE	Flk2 receptor protein-tyrosine-kinase cDNA.
XX	
KW	Human Flk2; receptor protein-tyrosine-kinase; primitive.
KW	hematopoietic cell; fetal liver kinase; ds.
XX	
OS	Homo sapiens.
XX	
PH	Key
FT	CDS
FT	Location/Qualifiers
FT	58..3039
FT	/tag= a
FT	/product= Flk2 receptor protein-tyrosine-kinase
FT	58..138
FT	/tag= b
FT	139..3036
FT	/tag= c
XX	
XX	WO9500554-A.
PN	
XX	
PD	05-JAN-1995.
XX	
PF	17-JUN-1994; 94WO-US06944.
XX	
PR	18-JUN-1993; 93US-0080244.
PR	21-JUN-1993; 93US-0081508.

PR 23-NOV-1993; 93US-0157490.  
PA (UYPR-) UNIV PRINCETON.  
XX Lemischka IR;  
XX WPI: 1995-052014/07.  
DR P-PSDB; AAR67816.  
XX  
PT Ligand for receptor protein tyrosine kinase - useful for the  
PT stimulation of primitive haematopoietic stem cells causing  
PT proliferation and/or differentiation  
XX  
PS Disclosure; Fig 1b; 131pp; English.  
XX  
CC The sequence corresponds to a cDNA encoding a human Flk2 (fetal  
CC liver kinase) receptor protein-tyrosine-kinase. Flk2 is expressed  
CC in primitive hematopoietic cells but not in mature hematopoietic  
CC cells. The gene product is useful in isolation of receptor  
CC ligands, which have applications in diagnosis of bone marrow  
CC disorders and in stimulating proliferation and/or differentiation  
CC of primitive hematopoietic stem cells.  
XX  
SQ Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;  
Query Match 100.0%; Score 3501; DB 16; Length 3501;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGAGCGGCATCCGAGGGCTGGCGCGCCCTGGGGACCCCGGGCTCCCGAGGCCATG 60  
DB 1 CGAGCGGCATCCGAGGGCTGGCGCGCCCTGGGGACCCCGGGCTCCCGAGGCCATG 60  
QY 61 CGGCGTGGCGGCGACGCGGCGACCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTG 120  
DB 61 CGGCGTGGCGGCGACGCGGCGACCGTGGCGCTGCTGCTGCTGCTGCTGCTGCTG 120  
QY 121 TTGSGACTATTACAAATCAAGATCTGCTGATCAAGTGTGTTTAAATCAATATAAG 180  
DB 121 TTGSGACTATTACAAATCAAGATCTGCTGATCAAGTGTGTTTAAATCAATATAAG 180  
QY 181 AACATGATTCATCAGTGGGAGTGCATCATATATCCCATGATCAGAAATCCCGGAA 240  
DB 181 AACATGATTCATCAGTGGGAGTGCATCATATATCCCATGATCAGAAATCCCGGAA 240  
QY 241 GACCTGGGTGTCGTTGAGACCCAGAGCTCAGGACAGTGTACGAGCTGCCGTGTG 300  
DB 241 GACCTGGGTGTCGTTGAGACCCAGAGCTCAGGACAGTGTACGAGCTGCCGTGTG 300  
QY 301 GAAGTGGATGTATCTGCTTCCATCAGTCAAGTGTGCTGCTGCTGCTGCTGCTGCTG 360  
DB 301 GAAGTGGATGTATCTGCTTCCATCAGTCAAGTGTGCTGCTGCTGCTGCTGCTGCTG 360  
QY 361 TCCTGCTCTGGGTCTTTAGACAGCTCCCTGAATGCCAGCCACATTTGATTTACAA 420  
DB 361 TCCTGCTCTGGGTCTTTAGACAGCTCCCTGAATGCCAGCCACATTTGATTTACAA 420  
QY 421 AACAGAGGAGTGTGTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC 480  
DB 421 AACAGAGGAGTGTGTTCCATGGTCAATTTGAAATGACAGAAACCCAGCTGGAGATAC 480  
QY 481 CTACTTTTATTCAGAGTGAAGTACCAATACACAATATGTTACAGTGAATTAAGA 540  
DB 481 CTACTTTTATTCAGAGTGAAGTACCAATACACAATATGTTACAGTGAATTAAGA 540  
QY 541 AATACCTGCTTTACATTAAGAGACCTTACTTTAGAAAATGGAACCCAGAGGCC 600  
DB 541 AATACCTGCTTTACATTAAGAGACCTTACTTTAGAAAATGGAACCCAGAGGCC 600  
QY 601 CTGCTCTGCATATCTGAGAGCGTTCAGAGCCGATCGTGAATGGGTGCTTGGGATTC 660  
DB 601 CTGCTCTGCATATCTGAGAGCGTTCAGAGCCGATCGTGAATGGGTGCTTGGGATTC 660

QY 661 CAGGGGAAAGCTGTAAAGAAAGAGTCCAGCTGTGTTTAAAGAGGAGAAAAGTGT 720  
DB 661 CAGGGGAAAGCTGTAAAGAAAGAGTCCAGCTGTGTTTAAAGAGGAGAAAAGTGT 720  
QY 721 CATGAATATTGAGGACGACATAGGTGCTGTCAGAAATGAACCTGGCAGGAATGC 780  
DB 721 CATGAATATTGAGGACGACATAGGTGCTGTCAGAAATGAACCTGGCAGGAATGC 780  
QY 781 ACCAGGCTGTTTCACAATAGATCTAAATCAAACTCTCAGACACATTTGCAATATT 840  
DB 781 ACCAGGCTGTTTCACAATAGATCTAAATCAAACTCTCAGACACATTTGCAATATT 840  
QY 841 CTTAAAGTAGGGAAACCCCTTATGGGATAGGTGCAAGCTGTTTCAATGCAATGATTC 900  
DB 841 CTTAAAGTAGGGAAACCCCTTATGGGATAGGTGCAAGCTGTTTCAATGCAATGATTC 900  
QY 901 GGGCTCACCTGGGAATTAGAAAACAAAGCACTCGAGAGGGCAACTTCTGAGATGAGT 960  
DB 901 GGGCTCACCTGGGAATTAGAAAACAAAGCACTCGAGAGGGCAACTTCTGAGATGAGT 960  
QY 961 ACCTATTCAACAAACAGAACTATGATACGATCTCTGCTTTTCTGCTTTTATCATCAGTGGCA 1020  
DB 961 ACCTATTCAACAAACAGAACTATGATACGATCTCTGCTTTTCTGCTTTTATCATCAGTGGCA 1020  
QY 1021 AGAAGCAGCACCGGATACACTTGTCTCTTCAAGCATCCCACTCAATCAGCTTGT 1080  
DB 1021 AGAAGCAGCACCGGATACACTTGTCTCTTCAAGCATCCCACTCAATCAGCTTGT 1080  
QY 1081 GTTACCCTAGTGGAGAGGATTTATAATGCTTACCAATTTCAAGTGAAGATTTGAAAT 1140  
DB 1081 GTTACCCTAGTGGAGAGGATTTATAATGCTTACCAATTTCAAGTGAAGATTTGAAAT 1140  
QY 1141 GACCAATATGAAGATTTTGTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200  
DB 1141 GACCAATATGAAGATTTTGTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200  
QY 1201 AGTGGACCTCTCTCGAAAATCATTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1260  
DB 1201 AGTGGACCTCTCTCGAAAATCATTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1260  
QY 1261 AGCATATCCAAGTTTTCGAATCATAGCACCAGGAGAAATATATATTTCCATGCAGAA 1320  
DB 1261 AGCATATCCAAGTTTTCGAATCATAGCACCAGGAGAAATATATATTTCCATGCAGAA 1320  
QY 1321 AATGATGATGCCCAATTTACCAAAATGTTTCCAGCTGAAATTAAGAAGAAACCTCAAGTG 1380  
DB 1321 AATGATGATGCCCAATTTACCAAAATGTTTCCAGCTGAAATTAAGAAGAAACCTCAAGTG 1380  
QY 1381 CTCGAGAGCATCGGCAAGTCAGGCTGCTGTTTCTCGGATGGATACCCATTTACCATCT 1440  
DB 1381 CTCGAGAGCATCGGCAAGTCAGGCTGCTGTTTCTCGGATGGATACCCATTTACCATCT 1440  
QY 1441 TGGACCTGGAAGAGTGTTCAGACAGTCTCCCAACTGCAGAGAGATCAGAGAGGA 1500  
DB 1441 TGGACCTGGAAGAGTGTTCAGACAGTCTCCCAACTGCAGAGAGATCAGAGAGGA 1500  
QY 1501 GTCTGGAATAGAAGGCTTAACAGAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
DB 1501 GTCTGGAATAGAAGGCTTAACAGAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
QY 1561 AACATGAGTGAAGCCATATAAGGTTTCTGCTCAAGTGTGTCGATACAAATTTCCCTTGGC 1620  
DB 1561 AACATGAGTGAAGCCATATAAGGTTTCTGCTCAAGTGTGTCGATACAAATTTCCCTTGGC 1620  
QY 1621 ACATCTTGTGAGAGCATCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCAACAGAAC 1680  
DB 1621 ACATCTTGTGAGAGCATCTTTTAAACTCTCCAGGCCCTTCCCTTTTCAATCAACAGAAC 1680  
QY 1681 ATCTCATCTTATGCAAAATTTGGTGTGCTCTCTCTTCAATTTGCTGTTTAAACCTGCTA 1740  
DB 1681 ATCTCATCTTATGCAAAATTTGGTGTGCTCTCTCTTCAATTTGCTGTTTAAACCTGCTA 1740  
QY 1741 ATTTGTCAAGTAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800



1741 ATTTGTCACAGTACAAAAGCAATTTAGTGTATGAAGCCAGCTACAGATGTTACAGGTG 1800  
1801 ACCGCTCCTCAGATTAAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860  
1801 ACCGCTCCTCAGATTAAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860  
1861 AAATGGGAGTTTCCCAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
1861 AAATGGGAGTTTCCCAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
1921 GGAAGATGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
1921 GGAAGATGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
1981 GCCGTCAAAATGCTGAAGAGAAAACGACAGACTCTGAAGAGAGGCACTCATGTCAGAA 2040  
1981 GCCGTCAAAATGCTGAAGAGAAAACGACAGACTCTGAAGAGAGGCACTCATGTCAGAA 2040  
2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAGATATTGTCAACCTGCTGGGGCGTGC 2100  
2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAGATATTGTCAACCTGCTGGGGCGTGC 2100  
2101 ACACGTGTCAGGACCAATTTACTTGTATTTTGAATCTGTTGCTATGTTGATCTTCTCAAC 2160  
2101 ACACGTGTCAGGACCAATTTACTTGTATTTTGAATCTGTTGCTATGTTGATCTTCTCAAC 2160  
2161 TATCTAAGAGTAAAGAGAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
2161 TATCTAAGAGTAAAGAGAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
2221 AATTTCAAGTTTACCACCTTTTCCAAATTCACATCAAAATTCACAGATGCTGTTCAAGA 2280  
2221 AATTTCAAGTTTACCACCTTTTCCAAATTCACATCAAAATTCACAGATGCTGTTCAAGA 2280  
2281 GAAGTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGGCTTCATCGGAAATTCATTTCAC 2340  
2281 GAAGTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGGCTTCATCGGAAATTCATTTCAC 2340  
2341 TCTGAAGATGAATTAAGATATGAATAACCAAAAGGCTGGAAGAGAGAGGAGTGAAT 2400  
2341 TCTGAAGATGAATTAAGATATGAATAACCAAAAGGCTGGAAGAGAGAGGAGTGAAT 2400  
2401 GTGCTTACATTTGAAGATCTCTTTGCTTTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
2401 GTGCTTACATTTGAAGATCTCTTTGCTTTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
2461 CTGGAATTTAAGTCTGTGTTTCACAGAGACCTGGCGCCAGGAGTGTCTGTCACCCAC 2520  
2461 CTGGAATTTAAGTCTGTGTTTCACAGAGACCTGGCGCCAGGAGTGTCTGTCACCCAC 2520  
2521 GGGAAAGTGGTGAAGATATGACATTTGGATTTGGCTCGAGATATCATGATGATTTCCAAAC 2580  
2521 GGGAAAGTGGTGAAGATATGACATTTGGATTTGGCTCGAGATATCATGATGATTTCCAAAC 2580  
2581 TATGTTGTCAGGGCAATGCCGCTGCTGCTGTAATGATGATGATGATGATGATGATGATGAT 2640  
2581 TATGTTGTCAGGGCAATGCCGCTGCTGCTGTAATGATGATGATGATGATGATGATGATGAT 2640  
2641 GAAGGCATCTACACCATTAAGAGTATGATGCTGCTATATGGAATATTACTGTGGGAATC 2700  
2641 GAAGGCATCTACACCATTAAGAGTATGATGCTGCTATATGGAATATTACTGTGGGAATC 2700  
2701 TTCTCAGTTGGTGAATCTTACCTGGCATTCGCGTTGATGCTAACTTCTACAAATG 2760  
2701 TTCTCAGTTGGTGAATCTTACCTGGCATTCGCGTTGATGCTAACTTCTACAAATG 2760  
2761 ATTCAAAATGGATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820  
2761 ATTCAAAATGGATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820  
2821 ATGCAATCTGCTGGCTTTTACACTCAAGGAACGCCATCTCCCTAATTTGACTTCG 2880

2821 ATGCAATCTGCTGGCTTTTGGACTCAAGAAACGGCATCTCTCCCTAATTTGACTTCG 2880  
2881 TTTTTAGGATGTCAGCTGGCAGATGCAGAGAAGCGATGTATCAGATGTGGATGGCGGT 2940  
2881 TTTTTAGGATGTCAGCTGGCAGATGCAGAGAAGCGATGTATCAGATGTGGATGGCGGT 2940  
2941 GTTTCGGAATGTCCTCACACCTACCAAAACAGCGGACCTTTTCAGCAGAGAGATGGATTG 3000  
2941 GTTTCGGAATGTCCTCACACCTACCAAAACAGCGGACCTTTTCAGCAGAGAGATGGATTG 3000  
3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060  
3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060  
3061 ACTTCATCCTCCGACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTATTTTCAT 3120  
3061 ACTTCATCCTCCGACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTATTTTCAT 3120  
3121 CACTAAAGAAAATCTATTATCAACTGCTGCTTCCAGCAGACTTTTCTCTAGAGCCCTCT 3180  
3121 CACTAAAGAAAATCTATTATCAACTGCTGCTTCCAGCAGACTTTTCTCTAGAGCCCTCT 3180  
3181 GCGTTTACTCTTGTGTTTCAAGGGACTTTTGTAAAAATCAATCATCTGTCACAAGGCG 3240  
3181 GCGTTTACTCTTGTGTTTCAAGGGACTTTTGTAAAAATCAATCATCTGTCACAAGGCG 3240  
3241 GAGGAGCTGATTAATGAATTTTATGGAGCATTCATCTGCATCCAAAGGCTTCTCAGGCG 3300  
3241 GAGGAGCTGATTAATGAATTTTATGGAGCATTCATCTGCATCCAAAGGCTTCTCAGGCG 3300  
3301 GCTTGAGTCAATTTGCTACCTGAAGTACAGTATATTCTTCTTAATACATAAACAAGC 3360  
3301 GCTTGAGTCAATTTGCTACCTGAAGTACAGTATATTCTTCTTAATACATAAACAAGC 3360  
3361 ATTTGCTAAGGAGAGCTAAATATGATTTTTTAAAGTCTATGTTTAAAAATATATGATA 3420  
3361 ATTTGCTAAGGAGAGCTAAATATGATTTTTTAAAGTCTATGTTTAAAAATATATGATA 3420  
3421 TTTTTCAGCTATTATGATATATTTTATGGTGGGAATAAATTTCTACTACAGAAAA 3480  
3421 TTTTTCAGCTATTATGATATATTTTATGGTGGGAATAAATTTCTACTACAGAAAA 3480  
3481 AAAAAAAAAAAAAAAAAAAAAA 3501  
3481 AAAAAAAAAAAAAAAAAAAAAA 3501

RESULT 2  
AAQ79069  
ID AAQ79069 standard; cdna; 3501 BP.  
XX AAQ79069;  
XX  
XX 04-JUL-1995 (first entry)  
XX Human flk-2 cdna.  
XX  
XX Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;  
KW hematopoiesis; stem cell; ds.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH CDS 58..3039  
FT /\*tag= a  
FT sig\_peptide 58..138  
FT /\*tag= b  
FT mat\_peptide 139..3036  
FT /\*tag= c  
XX  
XX US5367057-A.  
XX  
XX 22-NOV-1994.

XX 02-APR-1991; 91US-0679666.  
XX 02-APR-1991; 91US-0679666.  
PR 28-JUN-1991; 91US-0728913.  
PR 15-NOV-1991; 91US-0793065.  
PR 24-DEC-1991; 91US-0813593.  
PR 26-JUN-1992; 92US-0906397.  
PR 12-NOV-1992; 92US-0975049.  
PR 19-NOV-1992; 92US-0977451.  
PR 30-APR-1993; 93US-0055269.  
XX (UYPR-) UNIV PRINCETON.  
XX Lemischka IR;  
XX WPI; 1995-005894/01.  
DR P-PSDB; AAR67536.  
XX Murine flk-2 receptor protein tyrosine kinase - used to stimulate proliferation and/or stimulation of primitive mammalian haematopoietic stem cells in vitro or in vivo.  
XX Disclosure; Fig. 2A-1F; 69pp; English.  
XX cDNAs encoding receptor protein tyrosine-kinases, mouse fetal liver kinase-2 (flk-2), human flk-2 and mouse flk-1 are given in AAR79068-70, CC respectively, and the deduced amino acid sequences in AAR67535-37, CC respectively.  
XX Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;  
SQ  
Query Match 100.0%; Score 3501; DB 16; Length 3501;  
Best Local Similarity 100.0%; Pred. NO. 0;  
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGAGCGGCATCCGAGGCGTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60  
DB 1 CGAGCGGCATCCGAGGCGTGGCGCGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60  
QY 61 CGGCGGTTGGCGGCGACGCGGCGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
DB 61 CGGCGGTTGGCGGCGACGCGGCGCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
QY 121 TTGGGACTATTACAAATCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
DB 121 TTGGGACTATTACAAATCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
QY 181 AACAAATGATTCATAGTGGGAGTCAATCATATCCCATGCTATCAGAAATCCCGGAA 240  
DB 181 AACAAATGATTCATAGTGGGAGTCAATCATATCCCATGCTATCAGAAATCCCGGAA 240  
QY 241 GACCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
DB 241 GACCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
QY 301 GAAGTGGATGATTCATGCTTCCATCAGCTGCAAGTCTGCTGCTGCTGCTGCTGCTG 360  
DB 301 GAAGTGGATGATTCATGCTTCCATCAGCTGCAAGTCTGCTGCTGCTGCTGCTGCTG 360  
QY 361 TCCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 361 TCCCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
QY 421 AACAGAGGAGTGTGCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
DB 421 AACAGAGGAGTGTGCTTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
QY 481 CTACTTTTATTCAGAGTGAAGTACCAATTTACAAATTTGTTTACAGTGAATTAAGA 540  
DB 481 CTACTTTTATTCAGAGTGAAGTACCAATTTACAAATTTGTTTACAGTGAATTAAGA 540  
QY 541 AATACCTGCTTTACATTAAGAGACCTTACTTTAGAAAATGAAAACAGGACGCC 600

Db 541 AATACCTGCTTTACACATTAAAGACCTTACTTTAGAAAATGAAAACAGGACGCC 600  
QY 601 CTGGCTGCTATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGTGTCTTGGGATTC 660  
Db 601 CTGGCTGCTATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGTGTCTTGGGATTC 660  
QY 661 CAGGGGAAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTCTT 720  
Db 661 CAGGGGAAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTCTT 720  
QY 721 CATGAATTTATTTGGGACGACATAGTGTCTGCCAGAAATGAATGAGGAGGAAATGC 780  
Db 721 CATGAATTTATTTGGGACGACATAGTGTCTGCCAGAAATGAATGAGGAGGAAATGC 780  
QY 781 ACCAGGCTGTTACAAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCACAATTT 840  
Db 781 ACCAGGCTGTTCAAAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCACAATTT 840  
QY 841 CTTAAAGTAGGGGAAACCTTTATGGATAAGGTGCAAGCTGTTCTCATGTGAACCATGGATTC 900  
Db 841 CTTAAAGTAGGGGAAACCTTTATGGATAAGGTGCAAGCTGTTCTCATGTGAACCATGGATTC 900  
QY 901 GGCTCACCTGGGAATTAGAAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATGAGT 960  
Db 901 GGCTCACCTGGGAATTAGAAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATGAGT 960  
QY 961 ACCTATTCAAAACAGAACTATGATACGGATTTCTGTTGCTTTGATCATCATGAGTGC 1020  
Db 961 ACCTATTCAAAACAGAACTATGATACGGATTTCTGTTGCTTTGATCATCATGAGTGC 1020  
QY 1021 AGAAGCAGACCGGATACACTGTTCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080  
Db 1021 AGAAGCAGACCGGATACACTGTTCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080  
QY 1081 GTTACCATCGTAGAAGGGATTTATAAATGCTACCAATTTCAAGTGAAGATTTATGAAT 1140  
Db 1081 GTTACCATCGTAGAAGGGATTTATAAATGCTACCAATTTCAAGTGAAGATTTATGAAT 1140  
QY 1141 GACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCCTACCCCAAAATCAGATGT 1200  
Db 1141 GACCAATATGAAGATTTGTTTCTGTCAGGTTTAAAGCCTACCCCAAAATCAGATGT 1200  
QY 1201 ACCTGACCTTCTCTCGAAAATCATTTCTGAGCAAAAGGCTTGTGATAACGGATAC 1260  
Db 1201 ACCTGACCTTCTCTCGAAAATCATTTCTGAGCAAAAGGCTTGTGATAACGGATAC 1260  
QY 1261 AGCATATCCAAGTTTTCRAATCATAGCACCCAGGAGGAATATATATTCATGCAGAA 1320  
Db 1261 AGCATATCCAAGTTTTCRAATCATAGCACCCAGGAGGAATATATATTCATGCAGAA 1320  
QY 1321 AATGATGATGCCAATTTTACCAAAATGTTTCAAGTGAATTAAGAAGAAACCTCAAGTG 1380  
Db 1321 AATGATGATGCCAATTTTACCAAAATGTTTCAAGTGAATTAAGAAGAAACCTCAAGTG 1380  
QY 1381 CTCGCAAGAGCATCGGCAAGTCAAGGCTGCTGTTCTCGGATGGATACCATTTACCATCT 1440  
Db 1381 CTCGCAAGAGCATCGGCAAGTCAAGGCTGCTGTTCTCGGATGGATACCATTTACCATCT 1440  
QY 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACAAGAGGA 1500  
Db 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACAAGAGGA 1500  
QY 1501 GTCGTGAATAGAAAGCTTAACAGAAAAGTGTGGAGAGTGGGTGTCGAGCAGTACTCTA 1560  
Db 1501 GTCGTGAATAGAAAGCTTAACAGAAAAGTGTGGAGAGTGGGTGTCGAGCAGTACTCTA 1560  
QY 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGCTCAAGTCTGCTGCATACAAATTTCCCTTGGC 1620  
Db 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGCTCAAGTCTGCTGCATACAAATTTCCCTTGGC 1620  
QY 1621 ACATCTTTGAGACGATCCTTTTAAACTCTCCAGGCGCCCTTCCCTTTTCAAGAGCAAC 1680

Db 1621 ACATCTTGTGAGACGATCTCTTTAAACTCTCCAGGCCCTTCCCTTTCATCCAGACAAAC 1680  
Qy 1681 ATCTCATCTATCAACAATGGTGTGTTGCTCTCTCTCTCATTTGCTGCTTTTAAACCTGCTA 1740  
Db 1681 ATCTCATCTATCAACAATGGTGTGTTGCTCTCTCTCTCATTTGCTGCTTTTAAACCTGCTA 1740  
Qy 1741 ATTTGTACACAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Db 1741 ATTTGTACACAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Qy 1801 ACCGGCTCTCAGATATAGTACTCTTACCTTGATTTTTCAGAGATATGATATGATCTC 1860  
Db 1801 ACCGGCTCTCAGATATAGTACTCTTACCTTGATTTTTCAGAGATATGATATGATCTC 1860  
Qy 1861 AATGGGAGTTTCCAGAGAAATTTAGAGTTTGGAGGTACTAGGATCAGGTGCTTTT 1920  
Db 1861 AATGGGAGTTTCCAGAGAAATTTAGAGTTTGGAGGTACTAGGATCAGGTGCTTTT 1920  
Qy 1921 GGAAGAGTGTGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
Db 1921 GGAAGAGTGTGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
Qy 1981 GCCGTCAAATGCTGAAGAAAGAGACAGCTCTGAAAGAGAGGCACTCATGTFCAGAA 2040  
Db 1981 GCCGTCAAATGCTGAAGAAAGAGACAGCTCTGAAAGAGAGGCACTCATGTFCAGAA 2040  
Qy 2041 CTCAGATGATGCCCGAGCTGGAGGCCACAGATATTTGTAACCTCTGGGGCGTGC 2100  
Db 2041 CTCAGATGATGCCCGAGCTGGAGGCCACAGATATTTGTAACCTCTGGGGCGTGC 2100  
Qy 2101 ACATGTGAGACCAATTTACTTGATTTTGAATACTGTTGCTATGTTGATCTTCTCAAC 2160  
Db 2101 ACATGTGAGACCAATTTACTTGATTTTGAATACTGTTGCTATGTTGATCTTCTCAAC 2160  
Qy 2161 TATCTAAGAGTAAAGAGAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
Db 2161 TATCTAAGAGTAAAGAGAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
Qy 2221 AATTTTCAAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGT 2280  
Db 2221 AATTTTCAAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGT 2280  
Qy 2281 GAAGTTCAGATACACCCGAGCTCGATCAAAATCTCAGGCTTCATGGGAATTCATTTCAC 2340  
Db 2281 GAAGTTCAGATACACCCGAGCTCGATCAAAATCTCAGGCTTCATGGGAATTCATTTCAC 2340  
Qy 2341 TCTGAAGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 2400  
Db 2341 TCTGAAGATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 2400  
Qy 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
Db 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
Qy 2461 CTGGAATTTAAGTCTGTGTTACAGAGACCTGGCCGCGAGGAGCTGCTGTCCACCCAC 2520  
Db 2461 CTGGAATTTAAGTCTGTGTTACAGAGACCTGGCCGCGAGGAGCTGCTGTCCACCCAC 2520  
Qy 2521 GGAAGAGTGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTATTCACAC 2580  
Db 2521 GGAAGAGTGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTATTCACAC 2580  
Qy 2581 TATGTTGTGAGGGGCAATGCCCGTCTGCTGCTTAAATGGATGGGCCCCGGAAGCGCTGTTT 2640  
Db 2581 TATGTTGTGAGGGGCAATGCCCGTCTGCTGCTTAAATGGATGGGCCCCGGAAGCGCTGTTT 2640  
Qy 2641 GAAGGCATCTACACCATTAAGAGTATGCTGCTGCTATGATGATGATGATGATGATGATGATG 2700  
Db 2641 GAAGGCATCTACACCATTAAGAGTATGCTGCTGCTATGATGATGATGATGATGATGATGATG 2700  
Qy 2701 TTCTCAGTGTGTGAATCTTACCTTACCTGGGATTCGGGTTGATGCTTAACCTTCTACAACTG 2760  
Db 2701 TTCTCAGTGTGTGAATCTTACCTTACCTGGGATTCGGGTTGATGCTTAACCTTCTACAACTG 2760

Qy 2761 ATTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATTATA 2820  
Db 2761 ATTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATTATA 2820  
Qy 2821 ATGCAATCTGCTGGGCTTTTGGCTTTCAGTCAAGGAAACGCCCATCTCCCTAAATTTGACTTCG 2880  
Db 2821 ATGCAATCTGCTGGGCTTTTGGCTTTCAGTCAAGGAAACGCCCATCTCCCTAAATTTGACTTCG 2880  
Qy 2881 TTTTGTAGATGTACCTGTCAGATGCAAGAAGCGATGTATCAGATGTGGATGGCCGT 2940  
Db 2881 TTTTGTAGATGTACCTGTCAGATGCAAGAAGCGATGTATCAGATGTGGATGGCCGT 2940  
Qy 2941 GTTTCGGAATGTCCTCACACCTTACCAAAACAGCGACCTTTTCAGCAGAGAGATGGATTTCG 3000  
Db 2941 GTTTCGGAATGTCCTCACACCTTACCAAAACAGCGACCTTTTCAGCAGAGAGATGGATTTCG 3000  
Qy 3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG 3060  
Db 3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG 3060  
Qy 3061 ACTTCATCTCTCCACCTATCCCTACAGGCTGTAGATTACCAAAACAAGATTTAATTTTCAT 3120  
Db 3061 ACTTCATCTCTCCACCTATCCCTACAGGCTGTAGATTACCAAAACAAGATTTAATTTTCAT 3120  
Qy 3121 CACTAAAAGAAATCTATTATCAACTGCTGCTTCCACGACTTTTCTCTAGAGGCCGTCT 3180  
Db 3121 CACTAAAAGAAATCTATTATCAACTGCTGCTTCCACGACTTTTCTCTAGAGGCCGTCT 3180  
Qy 3181 CGGTTTACTCTTGTTTTCAAGGGGACTTTTGTAAAATCAAATCATCTGTCAACAAGCAG 3240  
Db 3181 CGGTTTACTCTTGTTTTCAAGGGGACTTTTGTAAAATCAAATCATCTGTCAACAAGCAG 3240  
Qy 3241 GAGGAGCTGATTAATGAATTTTATGGAGCATTTGATCTGCATCCAGGCCCTTCTCAGGCCG 3300  
Db 3241 GAGGAGCTGATTAATGAATTTTATGGAGCATTTGATCTGCATCCAGGCCCTTCTCAGGCCG 3300  
Qy 3301 GCTTCAGTGAATTTGTACCTGAAGTACAGTATTTCTTGTAAATACATAAACAAGC 3360  
Db 3301 GCTTCAGTGAATTTGTACCTGAAGTACAGTATTTCTTGTAAATACATAAACAAGC 3360  
Qy 3361 ATTTTGTCTAAGGAGAGCTAAATATGATTTTAAAGTCTATGTTTAAATATATATATAA 3420  
Db 3361 ATTTTGTCTAAGGAGAGCTAAATATGATTTTAAAGTCTATGTTTAAATATATATATAA 3420  
Qy 3421 TTTTTCAGCTATTTAGTATATTTTATGGTGGGAATATAAATTTCTTACTACAGAAAA 3480  
Db 3421 TTTTTCAGCTATTTAGTATATTTTATGGTGGGAATATAAATTTCTTACTACAGAAAA 3480  
Qy 3481 AAAAAA 3501  
Db 3481 AAAAAA 3501

## RESULT 3

ART38734

ID ART38734 standard: cDNA; 3501 BP.

XX AAT38734;

XX AC AAT38734;

XX XX 11-DEC-1996 (first entry)

XX Murine foetal liver kinase 2 cDNA.

DE Murine foetal liver kinase 2; flk-2; protein tyrosine kinase;

KW monoclonal; antibody; extracellular domain; receptor assay;

KW haematopoietic stem cell; ligand; stimulation; proliferation;

KW differentiation; treatment; anaemia; bone marrow damage;

KW cancer chemotherapy; radiation; ds.

XX Mus musculus.

XX OS Mus musculus.

XX FH Key

FH Location/Qualifiers

```

FT CDS 58...3039
FT sig_peptide /*tag= a
FT 58..138
FT mat_peptide /*tag= b
FT 139..3036
FT /*tag= c
XX
XX US5548065-A.
XX
XX 20-AUG-1996.
XX
XX 02-APR-1991; 91US-0679666.
XX
XX 19-NOV-1992; 92US-0977451.
XX 02-APR-1991; 91US-0679666.
XX 28-JUN-1991; 91US-0728913.
XX 15-NOV-1991; 91US-0793065.
XX 24-DEC-1991; 91US-0813593.
XX 26-JUN-1992; 92US-0906397.
XX 12-NOV-1992; 92US-0975049.
XX 30-APR-1993; 93US-0055269.
XX 31-OCT-1994; 94US-0252517.
XX
XX (UYPR-) UNIV PRINCETON.
XX
XX Lemischka IR;
XX
XX WPI; 1996-392678/39.
XX P-PSDB; AAR97419.
XX
XX Anti-fœtal liver kinase 2 (flk-2) antibodies - useful in assays,
XX for isolating haematopoietic stem cells expressing receptor and for
XX obtaining ligands
XX
XX Disclosure; Columns 39-48; 50pp; English.
XX
XX The present sequence encodes murine foetal liver kinase 2 (flk-2),
XX a protein tyrosine kinase. Isolated antibodies, pref. monoclonal,
XX raised against the extracellular portion of flk-2 can be used to
XX assay for flk receptors on the surface of primitive haematopoietic
XX stem cells, and to isolate positive cells. The antibodies can also
XX be used, as, or to obtain ligands, which stimulate the proliferation
XX and/or differentiation of stem cells. The ligands can be used, e.g.
XX for treating anaemia, or bone marrow damage resulting from cancer
XX chemotherapy, or radiation.
XX
XX Sequence 3501 BP; 1068 A; 709 C; 784 G; 940 T; 0 other;
XX
XX Query Match 100.0%; Score 3501; DB 17; Length 3501;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 CGAGGCGGCATCCGAGGGCTGGGCGGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60
Dd 1 CGAGGCGGCATCCGAGGGCTGGGCGGCCCTGGGGACCCCGGGCTCCGGAGGCCATG 60
QY 61 CGGCGCTTGGCGGCGGACCGGCGCCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 120
Dd 61 CGGCGCTTGGCGGCGGACCGGCGCCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 121 TTGCGGACTTACAAATCAAGATCTGCTGTGATCAAGTGCTGCTGCTGCTGCTGCTGCTG 180
Dd 121 TTGCGGACTTACAAATCAAGATCTGCTGTGATCAAGTGCTGCTGCTGCTGCTGCTGCTG 180
QY 181 ACAAATGATTCATCACTGGGAGTGCATCATATCCCATGCTGCTGCTGCTGCTGCTGCTG 240
Dd 181 ACAAATGATTCATCACTGGGAGTGCATCATATCCCATGCTGCTGCTGCTGCTGCTGCTG 240
QY 241 GACCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Dd 241 GACCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 301 GAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTG 360

```

Db 1381 CTCGAGAGCATCGGCAAGTCAGCGGTCCTGTTTCTCGGATGATACCCATTACCATCT 1440  
Qy 1441 TGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACAGAAGGA 1500  
Db 1441 TGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACAGAAGGA 1500  
Qy 1501 GTCTGGAATGAAGAAGGCTAACAGAAAAGTGTTCGACAGTGGTCTCGAGCAGTACTCTA 1560  
Db 1501 GTCTGGAATGAAGAAGGCTAACAGAAAAGTGTTCGACAGTGGTCTCGAGCAGTACTCTA 1560  
Qy 1561 AACATGAGTGAAGCCATGAAGAAGTGTTCGTCAGAGTGTGTGATACAAATTCCTTGGC 1620  
Db 1561 AACATGAGTGAAGCCATGAAGAAGTGTTCGTCAGAGTGTGTGATACAAATTCCTTGGC 1620  
Qy 1621 ACATCTGTGAGACGATCCTTTTAAACCTCCAGGCCCTTCCCTTTTCATCCACAGACAAC 1680  
Db 1621 ACATCTGTGAGACGATCCTTTTAAACCTCCAGGCCCTTCCCTTTTCATCCACAGACAAC 1680  
Qy 1681 ATCTCATCTATGAACAATTTGGTGTTCCTCTCTTCATTTGTCGTTTAAACCTTGCTA 1740  
Db 1681 ATCTCATCTATGAACAATTTGGTGTTCCTCTCTTCATTTGTCGTTTAAACCTTGCTA 1740  
Qy 1741 ATTTGTACAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGTG 1800  
Db 1741 ATTTGTACAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGTG 1800  
Qy 1801 ACCGCTCTCAGATAATGAGTACTCTACCTGATTTTCAGAGATATGAATATGATCTC 1860  
Db 1801 ACCGCTCTCAGATAATGAGTACTCTACCTGATTTTCAGAGATATGAATATGATCTC 1860  
Qy 1861 AAATGGGAGTTTCCAAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
Db 1861 AAATGGGAGTTTCCAAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
Qy 1921 GGAAGAAGTGAAGCAACAGCTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980  
Db 1921 GGAAGAAGTGAAGCAACAGCTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980  
Qy 1981 GCCGTCAAAATGCTGAAGAAAAAGCAGACAGCTCTGAAGAGAGGCACTCATGTCAAA 2040  
Db 1981 GCCGTCAAAATGCTGAAGAAAAAGCAGACAGCTCTGAAGAGAGGCACTCATGTCAAA 2040  
Qy 2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAGATATTTGAACCTCTGGGGCGTGC 2100  
Db 2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAGATATTTGAACCTCTGGGGCGTGC 2100  
Qy 2101 ACATGTCAGACCAATTTACTTGAATTTTGAATCTGCTATGCTATGCTTCTCAAC 2160  
Db 2101 ACATGTCAGACCAATTTACTTGAATTTTGAATCTGCTATGCTATGCTTCTCAAC 2160  
Qy 2161 TATCTAAGAAGTAAAGAAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
Db 2161 TATCTAAGAAGTAAAGAAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
Qy 2221 AATTTTCAGTTTACCCTTCAATCCATCCAAATTCAGCATCCCTGGTTCAAGA 2280  
Db 2221 AATTTTCAGTTTACCCTTCAATCCATCCAAATTCAGCATCCCTGGTTCAAGA 2280  
Qy 2281 GAAGTTCAGATACACCCGAGCTCGATCAAAATCTCAGGGCTTCATGGGAATTCATTCCAC 2340  
Db 2281 GAAGTTCAGATACACCCGAGCTCGATCAAAATCTCAGGGCTTCATGGGAATTCATTCCAC 2340  
Qy 2341 TCTGAAGATGAATTTGAATATGAACCAAAAGGCTGGAGAGAGAGAGGACTTGAAT 2400  
Db 2341 TCTGAAGATGAATTTGAATATGAACCAAAAGGCTGGAGAGAGAGAGGACTTGAAT 2400  
Qy 2401 GTGCTTACATTTGAAGATCTTCTTGTCTTCATATCAAGTTGCCAAGGAATGGAATTT 2460  
Db 2401 GTGCTTACATTTGAAGATCTTCTTGTCTTCATATCAAGTTGCCAAGGAATGGAATTT 2460  
Qy 2461 CTGGAATTTAAGTCTGTGTTTCAGAGACACTGGCCGCGCAGGAACGCTGTTGTCAACCCAC 2520  
Db 2461 CTGGAATTTAAGTCTGTGTTTCAGAGACACTGGCCGCGCAGGAACGCTGTTGTCAACCCAC 2520

RESULT 4  
AAT72117  
ID AAT72117 standard; cDNA; 3501 BP.

Qy 2521 GGGAAAGTGGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATCATGAGTATTCCAAC 2580  
Db 2521 GGGAAAGTGGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATCATGAGTATTCCAAC 2580  
Qy 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAAATGATGGCCCCGAAAGCCCTGTTT 2640  
Db 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAAATGATGGCCCCGAAAGCCCTGTTT 2640  
Qy 2641 GAAGGCATCTACACCATTAAGAGTGTGCTGGTCTATATGGATATATCTGTGGGAATC 2700  
Db 2641 GAAGGCATCTACACCATTAAGAGTGTGCTGGTCTATATGGATATATCTGTGGGAATC 2700  
Qy 2701 TTCTCACCTTGGTGTGAATCCTTTACCTGGCATTTCCGTTGATGCTAACTTCTACAACCTG 2760  
Db 2701 TTCTCACCTTGGTGTGAATCCTTTACCTGGCATTTCCGTTGATGCTAACTTCTACAACCTG 2760  
Qy 2761 ATTTCAAAATGGAATTTAAAATGATCAGCCATTTTATGCTACAGAGAATAATACATTTATA 2820  
Db 2761 ATTTCAAAATGGAATTTAAAATGATCAGCCATTTTATGCTACAGAGAATAATACATTTATA 2820  
Qy 2821 ATGCAATCCTGCTGGGCTTTTTCACCTCAAGGAAACGGCCATCTTCCCTAACTTTCAGCTTCG 2880  
Db 2821 ATGCAATCCTGCTGGGCTTTTTCACCTCAAGGAAACGGCCATCTTCCCTAACTTTCAGCTTCG 2880  
Qy 2881 TTTTTCAGATGTCAGCTGGCAGATGCAGAAAGCGATGTATCAGAAATGTGGATGGCCGT 2940  
Db 2881 TTTTTCAGATGTCAGCTGGCAGATGCAGAAAGCGATGTATCAGAAATGTGGATGGCCGT 2940  
Qy 2941 GTTTCGGAATGTCCTCACACCTTACCAGGAGGCGACCTTTTCAGCAGAGAGATGGATTG 3000  
Db 2941 GTTTCGGAATGTCCTCACACCTTACCAGGAGGCGACCTTTTCAGCAGAGAGATGGATTG 3000  
Qy 3001 GGGCTACTCTCTCCGAGGCTCAGGTGGAAGATTCGTAGAGAAACAATTTAGTTTAAAG 3060  
Db 3001 GGGCTACTCTCTCCGAGGCTCAGGTGGAAGATTCGTAGAGAAACAATTTAGTTTAAAG 3060  
Qy 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTTTCAT 3120  
Db 3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTTTCAT 3120  
Qy 3121 CACTAAAAGAAAATCTATTATCAACTGCTGCTTCCAGCAGACTTTTCTCTAGAGCCGCTCT 3180  
Db 3121 CACTAAAAGAAAATCTATTATCAACTGCTGCTTCCAGCAGACTTTTCTCTAGAGCCGCTCT 3180  
Qy 3181 GCGTTTACTCTGTTTTCAGAGGAGCTTTTGTAAAATCAAAATCATCTCTGTCACAGGCAG 3240  
Db 3181 GCGTTTACTCTGTTTTCAGAGGAGCTTTTGTAAAATCAAAATCATCTCTGTCACAGGCAG 3240  
Qy 3241 GAGGAGCTGATTAATCAACTTTTATGGAGCATTTGATCTGCATCCAAGCCCTTCTCAGGCCG 3300  
Db 3241 GAGGAGCTGATTAATCAACTTTTATGGAGCATTTGATCTGCATCCAAGCCCTTCTCAGGCCG 3300  
Qy 3301 GCTTCAGTCAATTTGCTACCTGAAAGTACAGTATTTTGTAAATACATAAAACAAGC 3360  
Db 3301 GCTTCAGTCAATTTGCTACCTGAAAGTACAGTATTTTGTAAATACATAAAACAAGC 3360  
Qy 3361 ATTTTCAGTAAAGAGAGCTAATGATTTTAAAGTCTATGTTTAAATAATATATGTA 3420  
Db 3361 ATTTTCAGTAAAGAGAGCTAATGATTTTAAAGTCTATGTTTAAATAATATATGTA 3420  
Qy 3421 TTTTTCAGCTATTTAGTGTATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAAA 3480  
Db 3421 TTTTTCAGCTATTTAGTGTATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAAA 3480  
Qy 3481 AAAAAAAAAAAAAAAAAAAAAA 3501  
Db 3481 AAAAAAAAAAAAAAAAAAAAAA 3501



1141 GACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCCTTACCACAAATCAGATGT 1200  
1141 GACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCCTTACCACAAATCAGATGT 1200  
1201 AGTGGACCTTCTCGAATAATCTTCTGAGCAAAAGGCTTGTGATAACGGATAC 1260  
1201 AGTGGACCTTCTCGAATAATCTTCTGAGCAAAAGGCTTGTGATAACGGATAC 1260  
1261 AGCATATCAAGTTTGTCAATCATAGCACACCGCAGGAGCAATATATCCATGCAGAA 1320  
1261 AGCATATCAAGTTTGTCAATCATAGCACACCGCAGGAGCAATATATCCATGCAGAA 1320  
1321 AATGATGATGCCAAATTTACCAAAATGTTACAGCTGAATATAGAAGAAACCTCAAGTG 1380  
1321 AATGATGATGCCAAATTTACCAAAATGTTACAGCTGAATATAGAAGAAACCTCAAGTG 1380  
1381 CTCGCAAGATCGGCAAGTCAGGGCTCCTGTTTCTCGATGGATACCCATFACCATCT 1440  
1381 CTCGCAAGATCGGCAAGTCAGGGCTCCTGTTTCTCGATGGATACCCATFACCATCT 1440  
1441 TGGACCTGGAAGAGTGTTCACACAAAGTCTCCCACTGCACAGAGAGATCACAAGGA 1500  
1441 TGGACCTGGAAGAGTGTTCACACAAAGTCTCCCACTGCACAGAGAGATCACAAGGA 1500  
1501 GTCTGGAATAGAAAGCTTAACAGAAAGTGTGAGACAGTGGGTGTCGAGCAGTACTCTA 1560  
1501 GTCTGGAATAGAAAGCTTAACAGAAAGTGTGAGACAGTGGGTGTCGAGCAGTACTCTA 1560  
1561 AACATGAGTGAAGCCATAAAAGGTTCTGGTCAAGTGTGTCATACAAATCCCTTGGC 1620  
1561 AACATGAGTGAAGCCATAAAAGGTTCTGGTCAAGTGTGTCATACAAATCCCTTGGC 1620  
1621 ACATCTGTGACAGCATCCTTTAACTCTCAGGCCCCCTTCCCTTTTCATCCAAAGACAC 1680  
1621 ACATCTGTGACAGCATCCTTTAACTCTCAGGCCCCCTTCCCTTTTCATCCAAAGACAC 1680  
1681 ATCTCATTTATGCAACAAATGGTGTGTCCTCTTTCATGTCGCTTTTAACTCTGCTA 1740  
1681 ATCTCATTTATGCAACAAATGGTGTGTCCTCTTTCATGTCGCTTTTAACTCTGCTA 1740  
1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800  
1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800  
1801 ACCGGCTCTCAGATAATCAGTACTCTACGTTGATTTAGGAGAAATATGAATATGATCTC 1860  
1801 ACCGGCTCTCAGATAATCAGTACTCTACGTTGATTTAGGAGAAATATGAATATGATCTC 1860  
1861 AAATGGGAGTTTCCAAAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
1861 AAATGGGAGTTTCCAAAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
1921 GGAAAGTGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
1921 GGAAAGTGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
1981 GCGGTCAAAATGCTGAAAGAAAGAACGACAGACGCTCTGAAGAGAGGCACTCATGTCAGAA 2040  
1981 GCGGTCAAAATGCTGAAAGAAAGAACGACAGACGCTCTGAAGAGAGGCACTCATGTCAGAA 2040  
2041 CTCAGATGATCACCAGCTGGGAAGCCACGAGAAATTTGTAACCTGCTGGGGGCTGC 2100  
2041 CTCAGATGATCACCAGCTGGGAAGCCACGAGAAATTTGTAACCTGCTGGGGGCTGC 2100  
2101 ACACGTGTCAGGACCAATTTACTTGTATTTTGAATACTGTTGATGCTGATCTCTCAAC 2160  
2101 ACACGTGTCAGGACCAATTTACTTGTATTTTGAATACTGTTGATGCTGATCTCTCAAC 2160  
2161 TATCTAAGAAAGTAAAGAAAGAAATTTTCAAGAGCTTGGACAGAGATTTTCAAGGAACAC 2220  
2161 TATCTAAGAAAGTAAAGAAAGAAATTTTCAAGAGCTTGGACAGAGATTTTCAAGGAACAC 2220  
2221 AATTTACGTTTTTACCCCACTTTCCAAATCACATCCAAATTCACAGATGCTGTTCAAGA 2280

1221 AATTTACGTTTTTACCCCACTTTCCAAATCACATCCAAATTCACAGCATGCTGTTCAAGA 2280  
2281 GAAGTTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCCAC 2340  
2281 GAAGTTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCCAC 2340  
2341 TCTGAAGATGAATTTGAATATGAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
2341 TCTGAAGATGAATTTGAATATGAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
2401 GTGCTTACATTTCAAGATCTTCTTGTGCTTTCATATCAAGTTCGCAAGGAATCGAATTT 2460  
2401 GTGCTTACATTTCAAGATCTTCTTGTGCTTTCATATCAAGTTCGCAAGGAATCGAATTT 2460  
2461 CTGGAATTTAAGTCGTGTTTCACAGAGACCTGGCCGCCAGGACGCTGTTGCACCCAC 2520  
2461 CTGGAATTTAAGTCGTGTTTCACAGAGACCTGGCCGCCAGGACGCTGTTGCACCCAC 2520  
2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTTCCAAAC 2580  
2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTTCCAAAC 2580  
2581 TATGTTGTGAGGGGCAATGCCGCTCTGCCCTGTAATAATGGATGGCCGCCGAAAGCCTGTTT 2640  
2581 TATGTTGTGAGGGGCAATGCCGCTCTGCCCTGTAATAATGGATGGCCGCCGAAAGCCTGTTT 2640  
2641 GAAGGCATCTACACCATTAAAGAGTGATGCTGGTGCATATGGAATATTTACTGTGGGAATC 2700  
2641 GAAGGCATCTACACCATTAAAGAGTGATGCTGGTGCATATGGAATATTTACTGTGGGAATC 2700  
2701 TTCTCACTTGGTGTGAATCCTTACCCTGGCATTTCCGGTTGATGCTACTTCTTACAACATG 2760  
2701 TTCTCACTTGGTGTGAATCCTTACCCTGGCATTTCCGGTTGATGCTACTTCTTACAACATG 2760  
2761 ATTCAAAATGGATTTAAAATGGATCAGCCATTTATGCTACAGAGAAATATACATTTATA 2820  
2761 ATTCAAAATGGATTTAAAATGGATCAGCCATTTATGCTACAGAGAAATATACATTTATA 2820  
2821 ATGCAATCTCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCTTCCCTAATTTGACTTCG 2880  
2821 ATGCAATCTCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCTTCCCTAATTTGACTTCG 2880  
2881 TTTTTCAGATGTCAGCTGGCAGATGCAGAGAAGCGATGTATCAGATGTGGATGGCCGT 2940  
2881 TTTTTCAGATGTCAGCTGGCAGATGCAGAGAAGCGATGTATCAGATGTGGATGGCCGT 2940  
2941 GTTTCGGAATGTCCTCACACCTTACCACAGGCGACCTTTCAGCAGAGATGGAATTTG 3000  
2941 GTTTCGGAATGTCCTCACACCTTACCACAGGCGACCTTTCAGCAGAGATGGAATTTG 3000  
3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060  
3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTAAAG 3060  
3061 ACTTCATCCCTCACACCTTACCACAGGCTGTAGATTTACCAAAACAAGATTAATTTTCAT 3120  
3061 ACTTCATCCCTCACACCTTACCACAGGCTGTAGATTTACCAAAACAAGATTAATTTTCAT 3120  
3121 CACTAAAGAAATCTTATTCACACTGCTTTCACAGAGACTTTTCTCTAGAAGCCGCTCT 3180  
3121 CACTAAAGAAATCTTATTCACACTGCTTTCACAGAGACTTTTCTCTAGAAGCCGCTCT 3180  
3181 GCGTTTACTCTGTTTTCAAAGGACTTTTGTAAAATCAATCATCTCTGTCACAAGGCAG 3240  
3181 GCGTTTACTCTGTTTTCAAAGGACTTTTGTAAAATCAATCATCTCTGTCACAAGGCAG 3240  
3241 GAGGAGCTGATTAATGAATTTTATGGAGCATTTGATCTGATCCTCAAGGCCCTTCTCAGGCCG 3300  
3241 GAGGAGCTGATTAATGAATTTTATGGAGCATTTGATCTGATCCTCAAGGCCCTTCTCAGGCCG 3300  
3301 GCTTGAGTGAATGTGTAAGTACAGTATATTTCTTCTTAATACATATAAACAAGC 3360  
3301 GCTTGAGTGAATGTGTAAGTACAGTATATTTCTTCTTAATACATATAAACAAGC 3360







Db 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTGGCTTTGTATCATCATAGTGGCA 1020  
Qy 1021 AGAAACGACACCGGATCTACACATGTTCCCTTCAAGCATCCAGTCAATCAGCTTTG 1080  
Db 1021 AGAAACGACACCGGATCTACACATGTTCCCTTCAAGCATCCAGTCAATCAGCTTTG 1080  
Qy 1081 GTTACCATCGTAGGAAAGGGATTTATAAATGCTACCAATTCAGTGAAGATTTAGAAAT 1140  
Db 1081 GTTACCATCGTAGGAAAGGGATTTATAAATGCTACCAATTCAGTGAAGATTTAGAAAT 1140  
Qy 1141 GACCAATATGAAGAGTTTGTGTTCTGTCAGGTTTAAAGCTACCCACAATCAGATGT 1200  
Db 1141 GACCAATATGAAGAGTTTGTGTTCTGTCAGGTTTAAAGCTACCCACAATCAGATGT 1200  
Qy 1201 ACGTGGACCTTCTCTCGAAATCATTTCTGTCAGCAAAAGGCTTTGATAACGGATAC 1260  
Db 1201 ACGTGGACCTTCTCTCGAAATCATTTCTGTCAGCAAAAGGCTTTGATAACGGATAC 1260  
Qy 1261 AGCATATCCAAGTTTTCATATGAACACAGCAGGAGAGATATATATTCATTCAGAGAA 1320  
Db 1261 AGCATATCCAAGTTTTCATATGAACACAGCAGGAGAGATATATATTCATTCAGAGAA 1320  
Qy 1321 AATGATGATGCCAATTTACCAAAATGTTACGCTGAATATAGAAGAAACCTCAAGTG 1380  
Db 1321 AATGATGATGCCAATTTACCAAAATGTTACGCTGAATATAGAAGAAACCTCAAGTG 1380  
Qy 1381 CTCGCAAGCATCGCAAGTCAGCGCTCTGTTCTCGGATGGATACCCATTACCATCT 1440  
Db 1381 CTCGCAAGCATCGCAAGTCAGCGCTCTGTTCTCGGATGGATACCCATTACCATCT 1440  
Qy 1441 TGGACCTTGGAAAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAAGGA 1500  
Db 1441 TGGACCTTGGAAAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAAGGA 1500  
Qy 1501 GTCTGGAATAGAAAGGCTTACAGAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
Db 1501 GTCTGGAATAGAAAGGCTTACAGAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
Qy 1561 AACATGAGTGAAGCCATAAAGAGTCTCTGCTCAAGTCTGTCATACAAATCCCTTGGC 1620  
Db 1561 AACATGAGTGAAGCCATAAAGAGTCTCTGCTCAAGTCTGTCATACAAATCCCTTGGC 1620  
Qy 1621 ACATCTTGTGAGACGATCCCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGACAAC 1680  
Db 1621 ACATCTTGTGAGACGATCCCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGACAAC 1680  
Qy 1681 ATCTCATCTATGACAAATGGTGTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
Db 1681 ATCTCATCTATGACAAATGGTGTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
Qy 1741 ATTTGTACAGATACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGTG 1800  
Db 1741 ATTTGTACAGATACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGTG 1800  
Qy 1801 ACCGGCTCTCAGATAATGAGTACTTCTACCTTGATTTACAGAGATATGAATATGATCTC 1860  
Db 1801 ACCGGCTCTCAGATAATGAGTACTTCTACCTTGATTTACAGAGATATGAATATGATCTC 1860  
Qy 1861 AAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTCTTTT 1920  
Db 1861 AAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTCTTTT 1920  
Qy 1921 GGAAGATGATGAAGCAAGCTTATGGAATTTAGCAATTTAGCAAGAGGCTCTCAATCCAGTT 1980  
Db 1921 GGAAGATGATGAAGCAAGCTTATGGAATTTAGCAATTTAGCAAGAGGCTCTCAATCCAGTT 1980  
Qy 1981 GCCGTCAAAATGCTCAAAAGAAAAGCACAGCTCTGAAAGAGAGGCACTCATGTACAGAA 2040  
Db 1981 GCCGTCAAAATGCTCAAAAGAAAAGCACAGCTCTGAAAGAGAGGCACTCATGTACAGAA 2040  
Qy 2041 CTCAGATGATGACCCAGCTGGAGGCAAGCAGAGATATTTGTAACCTCTGGGGCGTGC 2100  
Db 2041 CTCAGATGATGACCCAGCTGGAGGCAAGCAGAGATATTTGTAACCTCTGGGGCGTGC 2100

Qy 2101 ACATGTCAGGACCAATTTACTTGTGATTTTGAATACTGTTGATGCTATGCTTCTCAAC 2160  
Db 2101 ACATGTCAGGACCAATTTACTTGTGATTTTGAATACTGTTGATGCTATGCTTCTCAAC 2160  
Qy 2161 TATCTAAGAAGTAAAGAGAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
Db 2161 TATCTAAGAAGTAAAGAGAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
Qy 2221 AATTTCAAGTTTACCCACACTTTCCAATCAGATCCAAATTCAGCATGCTGTTCAAGA 2280  
Db 2221 AATTTCAAGTTTACCCACACTTTCCAATCAGATCCAAATTCAGCATGCTGTTCAAGA 2280  
Qy 2281 GAAAGTTCAGATACACCCGACCTCGGATCAAAATCTCAGGCTTCATGGGAAATTCATTTCC 2340  
Db 2281 GAAAGTTCAGATACACCCGACCTCGGATCAAAATCTCAGGCTTCATGGGAAATTCATTTCC 2340  
Qy 2341 TCTGAAGATGAAATTTGAATATGAAAACCAAAAGGCTGGAAGAGAGAGGACTTTGAAT 2400  
Db 2341 TCTGAAGATGAAATTTGAATATGAAAACCAAAAGGCTGGAAGAGAGAGGACTTTGAAT 2400  
Qy 2401 GTGCTTACATTTGAAGATCTCTTTGCTTTGATATCAAGTTGCCAAAGAAATGGAATTT 2460  
Db 2401 GTGCTTACATTTGAAGATCTCTTTGCTTTGATATCAAGTTGCCAAAGAAATGGAATTT 2460  
Qy 2461 CTGGAATTTAAGTCGTGTTTTCACAGAGACCTGGCCGCGAGAAACGCTGTTGCACCCAC 2520  
Db 2461 CTGGAATTTAAGTCGTGTTTTCACAGAGACCTGGCCGCGAGAAACGCTGTTGCACCCAC 2520  
Qy 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGTTCGAGATATCATGATGATTTCCAAC 2580  
Db 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTTGGTTCGAGATATCATGATGATTTCCAAC 2580  
Qy 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAAATGATGGCCCCGAGAAACGCTGTTT 2640  
Db 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAAATGATGGCCCCGAGAAACGCTGTTT 2640  
Qy 2641 GAAGCATCTACACCATTAAGAGTGTGCTGCTCATATGGAATATTTACTGTGGGAAATC 2700  
Db 2641 GAAGCATCTACACCATTAAGAGTGTGCTGCTCATATGGAATATTTACTGTGGGAAATC 2700  
Qy 2701 TTTCACTTGTGTGAATCTTACCTTGGCATTCGGTTGATGCTAACTTCTACAAACTG 2760  
Db 2701 TTTCACTTGTGTGAATCTTACCTTGGCATTCGGTTGATGCTAACTTCTACAAACTG 2760  
Qy 2761 ATTCAAAATGGAATTTAAAATGATCAGCCATTTTATGCTACAGAGAAAATATACATATA 2820  
Db 2761 ATTCAAAATGGAATTTAAAATGATCAGCCATTTTATGCTACAGAGAAAATATACATATA 2820  
Qy 2821 ATGCAATCTGCTGGGCTTTTACCTCAAGGAAACGGCCATCTTCCCTTAATTTGACTTCG 2880  
Db 2821 ATGCAATCTGCTGGGCTTTTACCTCAAGGAAACGGCCATCTTCCCTTAATTTGACTTCG 2880  
Qy 2881 TTTTATGAGTGTGAGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2940  
Db 2881 TTTTATGAGTGTGAGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2940  
Qy 2941 GTTTCGGAATGTCCTACACCTTACCAAAACAGGCGACCTTTTTCAGCAGAGATGATTTG 3000  
Db 2941 GTTTCGGAATGTCCTACACCTTACCAAAACAGGCGACCTTTTTCAGCAGAGATGATTTG 3000  
Qy 3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAAATTTAGTTTAAAGG 3060  
Db 3001 GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAAATTTAGTTTAAAGG 3060  
Qy 3061 ACTTCATCCCTCAGCTATCTCCCTTAACAGGCTGTAGATTTACCAAAACAGATTAATTTCA 3120  
Db 3061 ACTTCATCCCTCAGCTATCTCCCTTAACAGGCTGTAGATTTACCAAAACAGATTAATTTCA 3120  
Qy 3121 CACTAAAAGAAAATCTATTTATCACTGCTTCCAGCAGATTTTCTCTAGAGAGCGTCT 3180  
Db 3121 CACTAAAAGAAAATCTATTTATCACTGCTTCCAGCAGATTTTCTCTAGAGAGCGTCT 3180



Db 841 CTTAAAGTAGGGAAACCTTATGATAAGGTGCAAGCTGTTCATGTAACCATGGATTC 900  
Qy 901 GGGCTACCTGGGAAATAGAAAACAAAGCACTCAGAGGGCAACTACTTTGAGATGAGT 960  
Db 901 GGGCTACCTGGGAAATAGAAAACAAAGCACTCAGAGGGCAACTACTTTGAGATGAGT 960  
Qy 961 ACCTATTCAACAACCAACTATGATACGGATTCTGTTTGTCTTTGTATCATCATAGTGGCA 1020  
Db 961 ACCTATTCAACAACCAACTATGATACGGATTCTGTTTGTCTTTGTATCATCATAGTGGCA 1020  
Qy 1021 AGAAGCAGACCCGATACACTTGTCTCTTCAAGAGATCCAGTCAATCAGCTTTG 1080  
Db 1021 AGAAGCAGACCCGATACACTTGTCTCTTCAAGAGATCCAGTCAATCAGCTTTG 1080  
Qy 1081 GTTACCATCTAGAAAGGGATTTATAATGCTACCAATTCAGTGAAGATTTAGAAAT 1140  
Db 1081 GTTACCATCTAGAAAGGGATTTATAATGCTACCAATTCAGTGAAGATTTAGAAAT 1140  
Qy 1141 GACCAATATCAAGAGTTTCTGTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200  
Db 1141 GACCAATATCAAGAGTTTCTGTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200  
Qy 1201 ACCTGGACCTTCTCTCGAAATCATTTCTGTCAGCAAAAGGCTTGTATAACGGATAC 1260  
Db 1201 ACCTGGACCTTCTCTCGAAATCATTTCTGTCAGCAAAAGGCTTGTATAACGGATAC 1260  
Qy 1261 AGCATATCCAAAGTTTGCATATCAATGACACAGCCAGGAGAAATATATTTCCATGCAAG 1320  
Db 1261 AGCATATCCAAAGTTTGCATATCAATGACACAGCCAGGAGAAATATATTTCCATGCAAG 1320  
Qy 1321 AATGATGATCCCAATTTACCAAAATGTTTACGCTGATATATAAGAAAGAAACCTCAAGTG 1380  
Db 1321 AATGATGATCCCAATTTACCAAAATGTTTACGCTGATATATAAGAAAGAAACCTCAAGTG 1380  
Qy 1381 CTGCGAAGCATCGGCAAGTCAGGCGTCTGTTCTCGGATGATACCCATTAACCATCT 1440  
Db 1381 CTGCGAAGCATCGGCAAGTCAGGCGTCTGTTCTCGGATGATACCCATTAACCATCT 1440  
Qy 1441 TGGACCTGGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAAGGA 1500  
Db 1441 TGGACCTGGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAAGGA 1500  
Qy 1501 GTCTGGAATAGAAAGGCTAACAGAAAGTGTTCGACAGTGGTGTGACAGTACTCTA 1560  
Db 1501 GTCTGGAATAGAAAGGCTAACAGAAAGTGTTCGACAGTGGTGTGACAGTACTCTA 1560  
Qy 1561 AACATGAGTGAAGCATATAAGGGTTCCTGGTCAAGTGTGTCATACAATTCCTTTGGC 1620  
Db 1561 AACATGAGTGAAGCATATAAGGGTTCCTGGTCAAGTGTGTCATACAATTCCTTTGGC 1620  
Qy 1621 ACATCTTGTGAGAGATCCCTTTTAACTCTCCAGSCCCCTTCCCTTTATCCCAAGAAC 1680  
Db 1621 ACATCTTGTGAGAGATCCCTTTTAACTCTCCAGSCCCCTTCCCTTTATCCCAAGAAC 1680  
Qy 1681 ATCTCATCTATGCAACAATGGTGTGTCCTCTCTTCAATGTGCTTTTAAACCCCTGCTA 1740  
Db 1681 ATCTCATCTATGCAACAATGGTGTGTCCTCTCTTCAATGTGCTTTTAAACCCCTGCTA 1740  
Qy 1741 ATTTGTCACAGTACAAAAGCAATTTAGGTATCAAGCCAGCTACAGATGGTACAGGTG 1800  
Db 1741 ATTTGTCACAGTACAAAAGCAATTTAGGTATCAAGCCAGCTACAGATGGTACAGGTG 1800  
Qy 1801 ACCGGCTCCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860  
Db 1801 ACCGGCTCCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860  
Qy 1861 AAATGGAGTTTCCAAGAAAAATTTAGAGTTTGGGAAGGTACTPAGGATCAGGTGCTTTT 1920  
Db 1861 AAATGGAGTTTCCAAGAAAAATTTAGAGTTTGGGAAGGTACTPAGGATCAGGTGCTTTT 1920  
Qy 1921 GGAAGATGATACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
Db 1921 GGAAGATGATACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT 1980

Db 1921 GGAAGATGATACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
Qy 1981 GCCGTCAAAATGCTGAAGAAAAAGCAGACAGCTCTGAAGAGAGAGGCACTCATGTTCAGAA 2040  
Db 1981 GCCGTCAAAATGCTGAAGAAAAAGCAGACAGCTCTGAAGAGAGAGGCACTCATGTTCAGAA 2040  
Qy 2041 CTCGAAGTATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100  
Db 2041 CTCGAAGTATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2100  
Qy 2101 ACCTGTGAGGACCAATTTACTTTGATTTTGAATACTGTTGCTATGTTGATCTTCTCAAC 2160  
Db 2101 ACCTGTGAGGACCAATTTACTTTGATTTTGAATACTGTTGCTATGTTGATCTTCTCAAC 2160  
Qy 2161 TATCTAAGAAGTAAAGAGAAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
Db 2161 TATCTAAGAAGTAAAGAGAAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
Qy 2221 AATTTCACTTTTACCCCACTTTTCCAAATCAACATCCAAATTTCCAGCATGCTGTTCAAGA 2280  
Db 2221 AATTTCACTTTTACCCCACTTTTCCAAATCAACATCCAAATTTCCAGCATGCTGTTCAAGA 2280  
Qy 2281 GAAGTTCAAGATACACCCGACTCGGATCAAAATCTCAGGCTTTCATGGGAATTTCAITTCAC 2340  
Db 2281 GAAGTTCAAGATACACCCGACTCGGATCAAAATCTCAGGCTTTCATGGGAATTTCAITTCAC 2340  
Qy 2341 TCTGAAGATGAAATGGAATATGAATATGAAGGCTGGAGAGAGAGAGAGAGAGACTTTGAAT 2400  
Db 2341 TCTGAAGATGAAATGGAATATGAATATGAAGGCTGGAGAGAGAGAGAGAGACTTTGAAT 2400  
Qy 2401 GTGCTTACATTTGAAGATCTCTTGTCTGTCATATCAAGTTTGCACAAAGGATTTGAATTT 2460  
Db 2401 GTGCTTACATTTGAAGATCTCTTGTCTGTCATATCAAGTTTGCACAAAGGATTTGAATTT 2460  
Qy 2461 CTGGAATTTAAGTCTGTTTACAGAGACCTGGCCGAGAGAGTGTCTGTCACCCAC 2520  
Db 2461 CTGGAATTTAAGTCTGTTTACAGAGACCTGGCCGAGAGAGTGTCTGTCACCCAC 2520  
Qy 2521 GGGAAAGTGTGAAGATATGACATTTGGATTTGGTTCGAGATATCATGATGATTTCCAAC 2580  
Db 2521 GGGAAAGTGTGAAGATATGACATTTGGATTTGGTTCGAGATATCATGATGATTTCCAAC 2580  
Qy 2581 TATGTTGTCAGGCAATGCCGCTGCTGCTGTAATGATGATGATGATGATGATGATGATGAT 2640  
Db 2581 TATGTTGTCAGGCAATGCCGCTGCTGCTGTAATGATGATGATGATGATGATGATGATGAT 2640  
Qy 2641 GAAGCATCTACACCATTAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 2700  
Db 2641 GAAGCATCTACACCATTAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 2700  
Qy 2701 TTCTCACTTGTGTGAATCCTTTACCTGGCATTTCCGGTTGATGCTTAACCTTCTACAACATG 2760  
Db 2701 TTCTCACTTGTGTGAATCCTTTACCTGGCATTTCCGGTTGATGCTTAACCTTCTACAACATG 2760  
Qy 2761 ATTCAAAATGATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820  
Db 2761 ATTCAAAATGATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820  
Qy 2821 ATGCAATCTCTGCGCTTTTGAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2880  
Db 2821 ATGCAATCTCTGCGCTTTTGAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2880  
Qy 2881 TTTTAGGATGTCAGCTGGCAGATGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2940  
Db 2881 TTTTAGGATGTCAGCTGGCAGATGCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2940  
Qy 2941 GTTTCGGAATGTCTTCACACCTTACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3000  
Db 2941 GTTTCGGAATGTCTTCACACCTTACCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3000  
Qy 3001 GGGCTACTCTCTCCGAGGCTCAGTCAAGATGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3060  
Db 3001 GGGCTACTCTCTCCGAGGCTCAGTCAAGATGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3060



Db 601 CTGGTCTGCATATCTAGAGCGTTTCCAGAGCCGATCGTGGAAATGGGCTCTTGGGATTCA 660  
QY 661 CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGCTTCTTAAAGAGGAGAAAGTGCCTT 720  
Db 661 CAGGGGAAAGCTGTAAGAGAAAGTCCAGCTGCTTCTTAAAGAGGAGAAAGTGCCTT 720  
QY 721 CATGAATATTATGGGAGCGGACATAAGGTGCTGCCAGAAATGAACCTGGGAGGAAATGC 780  
Db 721 CATGAATATTATGGGAGCGGACATAAGGTGCTGCCAGAAATGAACCTGGGAGGAAATGC 780  
QY 781 ACAGGCTGTTTCAATATAGATCTAAATCAAACTCCTCAGACCACATTCGCCACAATATTT 840  
Db 781 ACCAGGCTGTTTCAATATAGATCTAAATCAAACTCCTCAGACCACATTCGCCACAATATTT 840  
QY 841 CTTAAAGTAGGGAAACCTTATGGATAGGTGCAAGCTGTTTCAATGTAACCATGGATTC 900  
Db 841 CTTAAAGTAGGGAAACCTTATGGATAGGTGCAAGCTGTTTCAATGTAACCATGGATTC 900  
QY 901 GGGCTCACCTGGGAATTAGAAAACAAAGCACTCAGGAGGGCAACTACTTTGAGATGAGT 960  
Db 901 GGGCTCACCTGGGAATTAGAAAACAAAGCACTCAGGAGGGCAACTACTTTGAGATGAGT 960  
QY 961 ACCTATTCAACAAAGCACTATGATACGGATTCTGTTGCTTTTGTATCATCAGTGCA 1020  
Db 961 ACCTATTCAACAAAGCACTATGATACGGATTCTGTTGCTTTTGTATCATCAGTGCA 1020  
QY 1021 AGAAGCAGACCGGATACACTTGTCTCTTCAAGCATCCAGTCATCAGCTTGG 1080  
Db 1021 AGAAGCAGACCGGATACACTTGTCTCTTCAAGCATCCAGTCATCAGCTTGG 1080  
QY 1081 GTTACCATTCTAGGAAAGGATTTATAATGCTACCAATTTCAAGTGAAGATTTATGAAT 1140  
Db 1081 GTTACCATTCTAGGAAAGGATTTATAATGCTACCAATTTCAAGTGAAGATTTATGAAT 1140  
QY 1141 GACCAATPATGAAGATTTTGTCTGTGAGGTTTAAAGCTTCCACCAATCAGATGT 1200  
Db 1141 GACCAATPATGAAGATTTTGTCTGTGAGGTTTAAAGCTTCCACCAATCAGATGT 1200  
QY 1201 ACGTGGACCTTCTCAGAAATCATTTCTGTGAGCAAAAGGCTTGTGATAACGGATAC 1260  
Db 1201 ACGTGGACCTTCTCAGAAATCATTTCTGTGAGCAAAAGGCTTGTGATAACGGATAC 1260  
QY 1261 AGCATATCCAAATTTTCCAAATCATPAAGCAGCAGGAGAAATATATATTCATCCAGAA 1320  
Db 1261 AGCATATCCAAATTTTCCAAATCATPAAGCAGCAGGAGAAATATATATTCATCCAGAA 1320  
QY 1321 AATGATGATGCCAAATTTACCAAATGTTACGCTGAATATAGAAGGAAACCTCAAGTG 1380  
Db 1321 AATGATGATGCCAAATTTACCAAATGTTACGCTGAATATAGAAGGAAACCTCAAGTG 1380  
QY 1381 CTGCGAAGCATCGGCAAGTCCAGGCTGCTGTTTCTCGGATGGATACCCATTTACCATCT 1440  
Db 1381 CTGCGAAGCATCGGCAAGTCCAGGCTGCTGTTTCTCGGATGGATACCCATTTACCATCT 1440  
QY 1441 TGGACCTGGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACAGAAGGA 1500  
Db 1441 TGGACCTGGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACAGAAGGA 1500  
QY 1501 GTCTGGAATAGAAAGGCTAACAGAAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
Db 1501 GTCTGGAATAGAAAGGCTAACAGAAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
QY 1561 AACATGATGAGCCATTAAGGGTTCCTGTGTAAGTGTCTGTGATCAATATTCCTTGGC 1620  
Db 1561 AACATGATGAGCCATTAAGGGTTCCTGTGTAAGTGTCTGTGATCAATATTCCTTGGC 1620  
QY 1621 ACATCTTGTGAGAGATCCCTTTAACTCTCCAGGCCCCCTTCCCTTTCATCCAGACAAC 1680  
Db 1621 ACATCTTGTGAGAGATCCCTTTAACTCTCCAGGCCCCCTTCCCTTTCATCCAGACAAC 1680  
QY 1681 ATCTCATCTATGCAACAATTTGGTGTGCTCCTCTTCTCATTTGTCGTTTTAACCCCTGCTA 1740  
Db 1681 ATCTCATCTATGCAACAATTTGGTGTGCTCCTCTTCTCATTTGTCGTTTTAACCCCTGCTA 1740

QY 1741 ATTTGTCCAAAGTACAAAAAGCAATTTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Db 1741 ATTTGTCCAAAGTACAAAAAGCAATTTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
QY 1801 ACCGGCTCCTCAGATAATAGTACTTCTACGTTTGATTTTTCAGAGAATATGAATATGATCTC 1860  
Db 1801 ACCGGCTCCTCAGATAATAGTACTTCTACGTTTGATTTTTCAGAGAATATGAATATGATCTC 1860  
QY 1861 AAATGGAGTTTCCAAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTCTTTT 1920  
Db 1861 AAATGGAGTTTCCAAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTCTTTT 1920  
QY 1921 GGAAGAGTGTGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCAGGTT 1980  
Db 1921 GGAAGAGTGTGAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCAGGTT 1980  
QY 1981 GCCCTCAAAATGCTGAAAGAAAAGACAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
Db 1981 GCCCTCAAAATGCTGAAAGAAAAGACAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
QY 2041 CTCAAGATGATGACCCAGCTGGGAAGCCACAGAGATATTGTGAACCTGCTGGGGCGTGC 2100  
Db 2041 CTCAAGATGATGACCCAGCTGGGAAGCCACAGAGATATTGTGAACCTGCTGGGGCGTGC 2100  
QY 2101 ACCTGTGAGCAACCAATTTACTTGTATTTTGAATACTGTGCTATGCTGATCTTCTCAAC 2160  
Db 2101 ACCTGTGAGCAACCAATTTACTTGTATTTTGAATACTGTGCTATGCTGATCTTCTCAAC 2160  
QY 2161 TATCTAAGAACTAAAGAGAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
Db 2161 TATCTAAGAACTAAAGAGAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
QY 2221 AATTTAGTTTTACCOCACCTTCCAAATCAGATCCAAATTCAGATGCTGCTGTTCAAGA 2280  
Db 2221 AATTTAGTTTTACCOCACCTTCCAAATCAGATCCAAATTCAGATGCTGCTGTTCAAGA 2280  
QY 2281 GAAGTTCAGATACACCCGGACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTCAC 2340  
Db 2281 GAAGTTCAGATACACCCGGACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTCAC 2340  
QY 2341 TCTGAAGATGAAATGAAATATGAAACCAAAAGCTTGGAAAGAGAGGAGGACTTGAAT 2400  
Db 2341 TCTGAAGATGAAATGAAATATGAAACCAAAAGCTTGGAAAGAGAGGAGGACTTGAAT 2400  
QY 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
Db 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
QY 2461 CTGGAATTTAAGTCTGTTTCACAGAGACCTGGCCGCCAGGAAAGCTTGTTCACCCAC 2520  
Db 2461 CTGGAATTTAAGTCTGTTTCACAGAGACCTGGCCGCCAGGAAAGCTTGTTCACCCAC 2520  
QY 2521 GGGAAAGTGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGATGATTCACAC 2580  
Db 2521 GGGAAAGTGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATGATGATTCACAC 2580  
QY 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAATGATGGCCCCCGAAAGCCCTGTTT 2640  
Db 2581 TATGTTGTGAGGGCAATGCCCGTCTGCTGTAAATGATGGCCCCCGAAAGCCCTGTTT 2640  
QY 2641 GAAGCATCTACACCATTAAGAGTATGCTGCTCATATGGAATATTACTGTGGGAATC 2700  
Db 2641 GAAGCATCTACACCATTAAGAGTATGCTGCTCATATGGAATATTACTGTGGGAATC 2700  
QY 2701 TTCTCAGTGTGTAATCTTACCTGCAATTCGGTTGATGCTTAACCTTCTACAACTG 2760  
Db 2701 TTCTCAGTGTGTAATCTTACCTGCAATTCGGTTGATGCTTAACCTTCTACAACTG 2760  
QY 2761 ATTCAAAATGGAATTTAAATGATCAGGCAATTTTATGCTACAGAGAANAATATACATTATA 2820  
Db 2761 ATTCAAAATGGAATTTAAATGATCAGGCAATTTTATGCTACAGAGAANAATATACATTATA 2820



Db 541 AATACCCCTGCTTTACACATTAAGAGACGCTTACTTTAGAAAAATGAAAAACGAGGACGCC 600  
Qy 601 CTGGTCTGCATATCTGAGAGCCTTCCAGAGCCGATCGTGAATGGTGTCTTTGCGGATTC 660  
Db 601 CTGGTCTGCATATCTGAGAGCCTTCCAGAGCCGATCGTGAATGGTGTCTTTGCGGATTC 660  
Qy 661 CAGGGGAAAGCTGTAAGAGAAAGTCCAGTCTGTTGTTTAAAGAGGAGGAAAAAGTCTT 720  
Db 661 CAGGGGAAAGCTGTAAGAGAAAGTCCAGTCTGTTGTTTAAAGAGGAGGAAAAAGTCTT 720  
Qy 721 CATGAATATTTGGCAGCGACATAGGTGCTGTGCAGAAATGAACCTGGGAGGGAATGC 780  
Db 721 CATGAATATTTGGCAGCGACATAGGTGCTGTGCAGAAATGAACCTGGGAGGGAATGC 780  
Qy 781 ACCAGGCTGTTCCACAAATAGATCTAAATCAAACTCTCAGACACACATTTGCCACAATTTATTT 840  
Db 781 ACCAGGCTGTTCCACAAATAGATCTAAATCAAACTCTCAGACACACATTTGCCACAATTTATTT 840  
Qy 841 CTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCATGGATTC 900  
Db 841 CTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCATGGATTC 900  
Qy 901 GGGCTCACCTGGGAATTAGAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATGAGT 960  
Db 901 GGGCTCACCTGGGAATTAGAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATGAGT 960  
Qy 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTGTGCTTTGTATCATCACTGGCA 1020  
Db 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTGTGCTTTGTATCATCACTGGCA 1020  
Qy 1021 AGAAACGACCGGATCTACACTTGTTCCTTCAAAAGCATCCAGTCAATCAGCTTTG 1080  
Db 1021 AGAAACGACCGGATCTACACTTGTTCCTTCAAAAGCATCCAGTCAATCAGCTTTG 1080  
Qy 1081 GTTACCATCTAGGAAAGGATTTATAATGCTACCAATTCAGTGAAGATTTGAAATTT 1140  
Db 1081 GTTACCATCTAGGAAAGGATTTATAATGCTACCAATTCAGTGAAGATTTGAAATTT 1140  
Qy 1141 GACCAATATGAAGAGTGTGTTTTCTGTGAGTTTTAAAGCTTAAAGCTACCCACAATCAGATGT 1200  
Db 1141 GACCAATATGAAGAGTGTGTTTTCTGTGAGTTTTAAAGCTTAAAGCTACCCACAATCAGATGT 1200  
Qy 1201 ACGTGGACCTTCTCTCGAAATCAATTCCTTGTGAGCAAAAGGGTCTTGATAACGGATAC 1260  
Db 1201 ACGTGGACCTTCTCTCGAAATCAATTCCTTGTGAGCAAAAGGGTCTTGATAACGGATAC 1260  
Qy 1261 AGCATATCCAGTTTGGCAATCAAGACAGCAGGAGGAGATATATATTCATCCAGAA 1320  
Db 1261 AGCATATCCAGTTTGGCAATCAAGACAGCAGGAGGAGATATATATTCATCCAGAA 1320  
Qy 1321 AATGATGATGCCAATTTACCAAAATGTTACGCTGAATATAAGAGAAACCTCAAGTG 1380  
Db 1321 AATGATGATGCCAATTTACCAAAATGTTACGCTGAATATAAGAGAAACCTCAAGTG 1380  
Qy 1381 CTCGAGAAGCATCGGCAAGTCAGCGCTCTGTTCTCGGATGGATACCCATACCATCT 1440  
Db 1381 CTCGAGAAGCATCGGCAAGTCAGCGCTCTGTTCTCGGATGGATACCCATACCATCT 1440  
Qy 1441 TGGACCTGGAGAAGTGTTCAGACAGTCTCCCAACTGCACAGAAGATCACAGAAGA 1500  
Db 1441 TGGACCTGGAGAAGTGTTCAGACAGTCTCCCAACTGCACAGAAGATCACAGAAGA 1500  
Qy 1501 GTCTGGAATAGAAAGGCTTAACAGAAAAGTGTGACAGTGGTGTCCAGCAGTACTCTA 1560  
Db 1501 GTCTGGAATAGAAAGGCTTAACAGAAAAGTGTGACAGTGGTGTCCAGCAGTACTCTA 1560  
Qy 1561 AACATGAGTGAAGGCATAAAAGGTTCTGTGCTCAAGTGTGTGCATACAAATTCCTTGGC 1620  
Db 1561 AACATGAGTGAAGGCATAAAAGGTTCTGTGCTCAAGTGTGTGCATACAAATTCCTTGGC 1620  
Qy 1621 ACATCTTGTGAGAGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAGACAAC 1680  
Db 1621 ACATCTTGTGAGAGATCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAGACAAC 1680

Qy 1681 ATCTCATTTCTATCAACAATTTGGTGTGTTCTCTCTCTTCAATGTCGTTTTTAACCCCTGCTA 1740  
Db 1681 ATCTCATTTCTATCAACAATTTGGTGTGTTCTCTCTCTTCAATGTCGTTTTTAACCCCTGCTA 1740  
Qy 1741 ATTTGTCACAAGTACAAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800  
Db 1741 ATTTGTCACAAGTACAAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800  
Qy 1801 ACCGGCTCTCAGATTAATGAGTACTTCTACGTGTGATTTTCAGAGAATATGATATGATCTC 1860  
Db 1801 ACCGGCTCTCAGATTAATGAGTACTTCTACGTGTGATTTTCAGAGAATATGATATGATCTC 1860  
Qy 1861 AATATGGAGCTTTCCAAAGAGAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT 1920  
Db 1861 AATATGGAGCTTTCCAAAGAGAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT 1920  
Qy 1921 GGAAAGTGTATGAACCAACAGCTTATGGAATTTAGCAAAAACAGAGTCTCAATCCAGGTT 1980  
Db 1921 GGAAAGTGTATGAACCAACAGCTTATGGAATTTAGCAAAAACAGAGTCTCAATCCAGGTT 1980  
Qy 1981 GCGGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTTCAGAA 2040  
Db 1981 GCGGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTTCAGAA 2040  
Qy 2041 CTCAGATGATACCCAGCTGGGAAGCCAGGAGAATTTGTGAACCTGCTGGGGGCGTGC 2100  
Db 2041 CTCAGATGATACCCAGCTGGGAAGCCAGGAGAATTTGTGAACCTGCTGGGGGCGTGC 2100  
Qy 2101 ACACGTCTCAGGACCAATTTACTTGTGATTTTGAATCTGTTGCTATGGTGTATCTTCAAC 2160  
Db 2101 ACACGTCTCAGGACCAATTTACTTGTGATTTTGAATCTGTTGCTATGGTGTATCTTCAAC 2160  
Qy 2161 TATCTAAGAAAGTAAAGAGAAAATTTTCACAGACTTTGGACAGAGATTTTCAAGGAACAC 2220  
Db 2161 TATCTAAGAAAGTAAAGAGAAAATTTTCACAGACTTTGGACAGAGATTTTCAAGGAACAC 2220  
Qy 2221 AATTTTCAGTTTTTACCCCACTTTCCCAATCACATCCAAATTTCCAGACTGCTGGTCAAGA 2280  
Db 2221 AATTTTCAGTTTTTACCCCACTTTCCCAATCACATCCAAATTTCCAGACTGCTGGTCAAGA 2280  
Qy 2281 GAAGTTCAGATACACCCGACTCGGATCAAACTCAGGGCTTCATGGGAATTCATTTCAAC 2340  
Db 2281 GAAGTTCAGATACACCCGACTCGGATCAAACTCAGGGCTTCATGGGAATTCATTTCAAC 2340  
Qy 2341 TCTGAAGATGAAATTTGAATATGAAAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
Db 2341 TCTGAAGATGAAATTTGAATATGAAAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
Qy 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTTGCTATCAAGTTCGCCAAGGAATGGAAATTT 2460  
Db 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTTGCTATCAAGTTCGCCAAGGAATGGAAATTT 2460  
Qy 2461 CTGGAAATTTAAGTCTGTGTTTCACAGAGACTTGGCCGCCAGGAACGCTGTGTCACCCAC 2520  
Db 2461 CTGGAAATTTAAGTCTGTGTTTCACAGAGACTTGGCCGCCAGGAACGCTGTGTCACCCAC 2520  
Qy 2521 GGGAAAGTGGTGAAGATATGTCGCTTTGGATTTGGCTCGAGATATCATGAGTGTATCCAAC 2580  
Db 2521 GGGAAAGTGGTGAAGATATGTCGCTTTGGATTTGGCTCGAGATATCATGAGTGTATCCAAC 2580  
Qy 2581 TATGTTGTCCAGGGCAATGCCCGTCTGCCCTGTAAAATGATGGCTGCCCCCGGAAACCTGTTT 2640  
Db 2581 TATGTTGTCCAGGGCAATGCCCGTCTGCCCTGTAAAATGATGGCTGCCCCCGGAAACCTGTTT 2640  
Qy 2641 GAAGGCATCTACACCATTAAGAGTGTGCTGCTATGGAATATTTACTGTGGGAAATC 2700  
Db 2641 GAAGGCATCTACACCATTAAGAGTGTGCTGCTATGGAATATTTACTGTGGGAAATC 2700  
Qy 2701 TTCTCACCTTGGTGTGAATCCTTTACCCCTGGCATTTCCCGTTGATGCTAACTTCTACAACCTG 2760  
Db 2701 TTCTCACCTTGGTGTGAATCCTTTACCCCTGGCATTTCCCGTTGATGCTAACTTCTACAACCTG 2760









Db 2581 TATGTTGTCAGGGCAATGCCGCTGCTGCTGTAATAATGGATGGCCGCCGAAAGCCTGTTT 2640  
 QY 2641 GAAGGCATCTACACATTAAGAGTGAATGCTGGTCAATATGAATAATTACTGTGGGAATC 2700  
 Db 2641 GAAGGCATCTACACATTAAGAGTGAATGCTGGTCAATATGAATAATTACTGTGGGAATC 2700  
 QY 2701 TTCTCACATTGGTGAATCCTTACCTCGCATTCGGTTGATGCTAACTTCTACAACTG 2760  
 Db 2701 TTCTCACATTGGTGAATCCTTACCTCGCATTCGGTTGATGCTAACTTCTACAACTG 2760  
 QY 2761 ATTCAAAATGGATTAAATGGATGAGCAATTTATGCTACAGAGAATAATACATTATA 2820  
 Db 2761 ATTCAAAATGGATTAAATGGATGAGCAATTTATGCTACAGAGAATAATACATTATA 2820  
 QY 2821 ATGCAATCCTGCTGGGCTTTGACTCAAGGAACGGCCATCTTCCCTAAATTGACTTCG 2880  
 Db 2821 ATGCAATCCTGCTGGGCTTTGACTCAAGGAACGGCCATCTTCCCTAAATTGACTTCG 2880  
 QY 2881 TTTTATAGATGTCAGCTGCGAGATGCAGAAGAGCGATGTATCAGAATGTGGATGCCGT 2940  
 Db 2881 TTTTATAGATGTCAGCTGCGAGATGCAGAAGAGCGATGTATCAGAATGTGGATGCCGT 2940  
 QY 2941 GTTTCGGAATGTCCTCACACCTTACCAAAACAGCGACCTTTTCCAGCAGAGATGATTG 3000  
 Db 2941 GTTTCGGAATGTCCTCACACCTTACCAAAACAGCGACCTTTTCCAGCAGAGATGATTG 3000  
 QY 3001 GGGCTACTCTCTCCGAGCTCAGTGCAGAGATTGCTAGAGGAACAATTTAGTTTAAAG 3060  
 Db 3001 GGGCTACTCTCTCCGAGCTCAGTGCAGAGATTGCTAGAGGAACAATTTAGTTTAAAG 3060  
 QY 3061 ACTTCATCCCTCCACCTATCCCTAAACAGGCTGTAGATTACCAAAACAGATTAAATTAT 3120  
 Db 3061 ACTTCATCCCTCCACCTATCCCTAAACAGGCTGTAGATTACCAAAACAGATTAAATTAT 3120  
 QY 3121 CACTAAAAGAAAATCTATTATCAACTGCTGCTTCCAGCAGACTTTTCTAGAGCGCTCT 3180  
 Db 3121 CACTAAAAGAAAATCTATTATCAACTGCTGCTTCCAGCAGACTTTTCTAGAGCGCTCT 3180  
 QY 3181 GCGTTTACTCTGTTTCAAGGGACTTTTGTAAATCAATCAATCCTCTCAAGGCAG 3240  
 Db 3181 GCGTTTACTCTGTTTCAAGGGACTTTTGTAAATCAATCAATCCTCTCAAGGCAG 3240  
 QY 3241 GAGGAGCTGATAATGAACCTTATTGGAGCATGATCTGCATCCAAAGCGCTTCTCAGCGCG 3300  
 Db 3241 GAGGAGCTGATAATGAACCTTATTGGAGCATGATCTGCATCCAAAGCGCTTCTCAGCGCG 3300  
 QY 3301 GCTTGAGTGAATGTGTACCTGAAGTACAGTATATTCTTGTAAATACATAAAACAAGC 3360  
 Db 3301 GCTTGAGTGAATGTGTACCTGAAGTACAGTATATTCTTGTAAATACATAAAACAAGC 3360  
 QY 3361 ATTTTCTAGGAGAAGCTAATATGATTTTAAAGTCTATGTTTAAATATATCTAAA 3420  
 Db 3361 ATTTTCTAGGAGAAGCTAATATGATTTTAAAGTCTATGTTTAAATATATCTAAA 3420  
 QY 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGGGAATAAAATTTCTACTACAGA 3476  
 Db 3421 TTTTTCAGCTATTAGTGATATATTTTATGGTGGGAATAAAATTTCTACTACAGA 3476

## RESULT 10

AAT00802

ID AAT00802 standard; DNA; 3475 BP.

XX AC

XX AC

XX AC

DT 19-MAR-1996 (first entry)

XX Human Flk2/flt3 tyrosine kinase receptor gene.

DE

KW Tyrosine kinase receptor; flk2; flt3; agonist; monoclonal antibody;

KW haematopoiesis; hypoplasia; anaemia; thrombocytopenia; stem cell; ss.

XX

OS Homo sapiens.

XX Key Location/Qualifiers  
 PH CDS 58..3039  
 FT /tag= a  
 FT /product= flk2/flt3\_tyrosine\_kinase\_receptor  
 XX W09527062-A1.  
 XX 12-OCT-1995.  
 XX 23-MAR-1995; 95WO-US03718.  
 XX 04-APR-1994; 94US-0222299.  
 XX (GETH ) GENENTECH INC.  
 XX Bennett BD, Broz SD, Matthews W, Zeigler FC;  
 XX P-PSDB; AAR81869.  
 XX WPI; 1995-358636/46.  
 XX Agonist antibody which binds to flk2-flt3 tyrosine kinase receptor  
 XX - enhances proliferation of haematopoietic stem cells, in the  
 XX treatment of hypoplasia, anaemia, etc.  
 XX Disclosure; Page 41-44; 59pp; English.  
 XX DNA (AAT00801) coding for murine stem cell tyrosine kinase receptor flk2  
 XX (also called flt3) (AAR81868) was obtd. by RT-PCR amplification of RNA  
 XX isolated from mid-gestation mouse foetal livers using primers based  
 XX on the murine flt3 sequence, and subcloning of the product into pRK5.1.  
 XX An flk2/flt3 extracellular domain-IgG1 Fc fusion protein was  
 XX constructed and used to raise agonist antibodies able to bind to, and  
 XX activate, flk2/flt3. The human flk2/flt3 receptor gene (AAT00802) and  
 XX amino acid sequence (AAR81869) are also given.  
 XX SQ Sequence 3475 BP; 1042 A; 709 C; 784 G; 940 T; 0 other;  
 Query Match 99.1%; Score 3470.2; DB 16; Length 3475;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 3472; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CGAGCGCGCATCGAGGGCTGGCCGCGCCCTGGGGGACCCGGGCTCCGGAGGCCATG 60  
 Db 1 CGAGCGCGCATCGAGGGCTGGCCGCGCCCTGGGGGACCCGGGCTCCGGAGGCCATG 60  
 QY 61 CCGCGCTTGGCGCGCACCGCGGCACCGTCCCTGCTGCTGTTTCTGCAATGATA 120  
 Db 61 CCGCGCTTGGCGCGCACCGCGGCACCGTCCCTGCTGCTGTTTCTGCAATGATA 120  
 QY 121 TTTGGGACTATTACAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAAG 180  
 Db 121 TTTGGGACTATTACAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCAAG 180  
 QY 181 AACAAATGATTCAATCAGTGGGAAAGTATCATATATATATATATATATATATATAT 240  
 Db 181 AACAAATGATTCAATCAGTGGGAAAGTATCATATATATATATATATATATATATAT 240  
 QY 241 GACCTCGGTGTGCGTTGAGACCCAGCTCAGGACAGTGTACGAGCTGCCGCTGTG 300  
 Db 241 GACCTCGGTGTGCGTTGAGACCCAGCTCAGGACAGTGTACGAGCTGCCGCTGTG 300  
 QY 301 GAAGTGGATGATCTGCTTCCATCACACTGCAAGTGTGTTTAAATCAATCAAG 360  
 Db 301 GAAGTGGATGATCTGCTTCCATCACACTGCAAGTGTGTTTAAATCAATCAAG 360  
 QY 361 TCCTGTCTCTGGGTCTTTAAAGCAGAGCTCCCTGAAATGCCAGGCACATTGTTTACAA 420  
 Db 361 TCCTGTCTCTGGGTCTTTAAAGCAGAGCTCCCTGAAATGCCAGGCACATTGTTTACAA 420  
 QY 421 AACAGAGAGTGTGTTTCCATGTTTGAATAATGACAGAAACCCCAAGCTGGAGATAC 480  
 Db 421 AACAGAGAGTGTGTTTCCATGTTTGAATAATGACAGAAACCCCAAGCTGGAGATAC 480





Db 601 TCACAGGGGAAAGCTGTAAAGAAAGAGTCCAGCTGTGTGTAAAGAGGAGAAAAAGTG 660  
Qy 718 CTTTCATGAATATTTTGGGACGACATAAGGTGCTGTGCCAGAAATGAACCTGGGAGGAA 777  
Db 661 CTTTCATGAATATTTTGGGACGACATAAGGTGCTGTGCCAGAAATGAACCTGGGAGGAA 720  
Qy 778 TGCACAGGTGTTTCACAATAGATCTAAATCAAACTCCTCAGACACATTCGCCACAATTA 837  
Db 721 TGCACAGGTGTTTCACAATAGATCTAAATCAAACTCCTCAGACACATTCGCCACAATTA 780  
Qy 838 TTTCTTAAAGTAGGGGAACCTTATGATATAGGTGCAAAAGCTGTTCATGTGAACCATGGA 897  
Db 781 TTTCTTAAAGTAGGGGAACCTTATGATATAGGTGCAAAAGCTGTTCATGTGAACCATGGA 840  
Qy 898 TTCGGGCTCAGCTGGGAATTAGAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATG 957  
Db 841 TTCGGGCTCAGCTGGGAATTAGAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATG 900  
Qy 958 AGTACCTATTCAACAAACAGAACTATGATAGGATCTCTGTTTGTGCTTTGTATCATCAGTG 1017  
Db 901 AGTACCTATTCAACAAACAGAACTATGATAGGATCTCTGTTTGTGCTTTGTATCATCAGTG 960  
Qy 1018 GCAGAAACGACCCGGATACACTTGTTCCTCTTCAAGAGATCCAGTCAATCAGCT 1077  
Db 961 GCAGAAACGACCCGGATACACTTGTTCCTCTTCAAGAGATCCAGTCAATCAGCT 1020  
Qy 1078 TTGGTTACCATCGTAGGAAGGGATTTATAATCTACCATTCAAGTGAAGATTTACAA 1137  
Db 1021 TTGGTTACCATCGTAGGAAGGGATTTATAATCTACCATTCAAGTGAAGATTTACAA 1080  
Qy 1138 ATTGACCAATATGAAGAGTTTGTGTTTCTGTCAAGGTTTAAAGCTTACCACCAATCAGA 1197  
Db 1081 ATTGACCAATATGAAGAGTTTGTGTTTCTGTCAAGGTTTAAAGCTTACCACCAATCAGA 1140  
Qy 1198 TGTACGTGGACCTTCTCTCGAAATATCTTCTTGTGAGCAAAAGGCTTTGATTAACGGA 1257  
Db 1141 TGTACGTGGACCTTCTCTCGAAATATCTTCTTGTGAGCAAAAGGCTTTGATTAACGGA 1200  
Qy 1258 TACAGCATATCCAAAGTTTGCATATCAAGACACAGCCAGGAGATATATATTTCCATGCA 1317  
Db 1201 TACAGCATATCCAAAGTTTGCATATCAAGACACAGCCAGGAGATATATATTTCCATGCA 1260  
Qy 1318 GAAATGATGATGCCCAATTTACCAAAATGTTACGCTGAATATGAAGAGAAACCTCAA 1377  
Db 1261 GAAATGATGATGCCCAATTTACCAAAATGTTACGCTGAATATGAAGAGAAACCTCAA 1320  
Qy 1378 GTGCTCGAGAAGCATCGGCAAGTCAGCGCTCTGTTTCTCGGATGGATACCCATTACCA 1437  
Db 1321 GTGCTCGAGAAGCATCGGCAAGTCAGCGCTCTGTTTCTCGGATGGATACCCATTACCA 1380  
Qy 1438 TCTTGGACCTGGAAGAAGTGTTCAGACAAAGTCTCCCAACTCCAGAGAGATCACAGAA 1497  
Db 1381 TCTTGGACCTGGAAGAAGTGTTCAGACAAAGTCTCCCAACTCCAGAGAGATCACAGAA 1440  
Qy 1498 GGAGTCTGGAATACAAAGGCTAACAGAAAAGTGTGTCAGAGTGGGTGTCAGAGTACT 1557  
Db 1441 GGAGTCTGGAATACAAAGGCTAACAGAAAAGTGTGTCAGAGTGGGTGTCAGAGTACT 1500  
Qy 1558 CTAACATGATGAGGAGCCATAAAAGGTTCTGTCGAGTCTGTCATACAAATTCCTCT 1617  
Db 1501 CTAACATGATGAGGAGCCATAAAAGGTTCTGTCGAGTCTGTCATACAAATTCCTCT 1560  
Qy 1618 GGCACATCTGTGAGACGATCCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAGAC 1677  
Db 1561 GGCACATCTGTGAGACGATCCTTTTAACTCTCCAGGCCCTTCCCTTTTCATCCAGAC 1620  
Qy 1678 AACATCTCATTTATGCAACAATTTGGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1737  
Db 1621 AACATCTCATTTATGCAACAATTTGGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680  
Qy 1738 CTAATTTGTACAAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGATACAG 1797

Db 1691 CTAATTTGTACAAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGATACAG 1740  
Qy 1798 GTGACCGGCTCCTCAGATAAATGAGTACTTCTACGTTGATTTTTCAGAGAAATATGAATATGAT 1857  
Db 1741 GTGACCGGATCCTCAGATAAATGAGTACTTCTACGTTGATTTTTCAGAGAAATATGAATATGAT 1800  
Qy 1858 CTCAAATGGGAGTTTCCAAAGGAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCT 1917  
Db 1801 CTCAAATGGGAGTTTCCAAAGGAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCT 1860  
Qy 1918 TTTGGAAGTATGATCAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG 1977  
Db 1861 TTTGGAAGTATGATCAACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG 1920  
Qy 1978 GTTGCGCTCAAAATCCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGACACTCATGTCA 2037  
Db 1921 GTTACCGTCAAAATGCTCAAGAAAAGCAGACAGCTCTGAAAGAGAGGACACTCATGTCA 1980  
Qy 2038 GAACTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATTTGTGAACCTGCTGGGGCG 2097  
Db 1981 GAACTCAAGATGATGACCCAGCTGGGAAGCCACGAGAAATTTGTGAACCTGCTGGGGCG 2040  
Qy 2098 TGCACACTGTGAGGACCAATTTTACTTTGATTTTGAATACTGTTGCTATGCTGATCTCTC 2157  
Db 2041 TGCACACTGTGAGGACCAATTTTACTTTGATTTTGAATACTGTTGCTATGCTGATCTCTC 2100  
Qy 2158 AACATCTTAAGAAGTAAAGAGAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAA 2217  
Db 2101 AACATCTTAAGAAGTAAAGAGAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAA 2160  
Qy 2218 CACATTTTCAAGTTTACCCCACTTTTCAATCACATCCAAATTTCCAGCATGCTGCTGTTCA 2277  
Db 2161 CACATTTTCAAGTTTACCCCACTTTTCAATCACATCCAAATTTCCAGCATGCTGCTGTTCA 2220  
Qy 2278 AGAGAAGTTCAGATACACCCGACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTT 2337  
Db 2221 AGAGAAGTTCAGATACACCCGACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTT 2280  
Qy 2338 CACTCTGAAGATGAAATTCGAATATCAAAACCAAAAGGCTGGAAGAGAGGAGGACTTG 2397  
Db 2281 CACTCTGAAGATGAAATTCGAATATCAAAACCAAAAGGCTGGAAGAGAGGAGGACTTG 2340  
Qy 2398 AATGCTTACATTTGAAGATCTTCTTGTGCTTTCATATCAAGTTGCCAAGGAATGGAA 2457  
Db 2341 AATGCTTACATTTGAAGATCTTCTTGTGCTTTCATATCAAGTTGCCAAGGAATGGAA 2400  
Qy 2458 TTTCTGGAATTTAAGTCGTGTTTTCAGAGACCTGCGCCGAGGAACTGCTGTTGTCACC 2517  
Db 2401 TTTCTGGAATTTAAGTCGTGTTTTCAGAGACCTGCGCCGAGGAACTGCTGTTGTCACC 2460  
Qy 2518 CACGGAAAGTGGTGAAGATATGTCACATTTGGATTGGCTCGAGATATCATGATGATTC 2577  
Db 2461 CACGGAAAGTGGTGAAGATATGTCACATTTGGATTGGCTCGAGATATCATGATGATTC 2520  
Qy 2578 AACTATGTTGTGAGGGCAATGCCGCTCTGCTGCTGTAATGATGGCTGCGCCGAGGAACTG 2637  
Db 2521 AACTATGTTGTGAGGGCAATGCCGCTCTGCTGCTGTAATGATGGCTGCGCCGAGGAACTG 2580  
Qy 2638 TTTGAGGACATCTACACCATTAAGAGTATGTCGTGCTATGGAATATTTACTGTCGGAA 2697  
Db 2581 TTTGAGGACATCTACACCATTAAGAGTATGTCGTGCTATGGAATATTTACTGTCGGAA 2640  
Qy 2698 ATCTTCTCATTGCTGTAATCCTTTACCTGCAATTTCCGTTGATGCTCACTTCTACAAA 2757  
Db 2641 ATCTTCTCATTGCTGTAATCCTTTACCTGCAATTTCCGTTGATGCTCACTTCTACAAA 2700  
Qy 2758 CTGATTTCAAAATGGAATTTAAATGATGAGCAATTTATGCTACAGAAATATACATTT 2817  
Db 2701 CTGATTTCAAAATGGAATTTAAATGATGAGCAATTTATGCTACAGAAATATACATTT 2760  
Qy 2818 ATAATGCAATCCTGCTGGGCTTTTGAACCTCAAGGAAACGCCATCTTCCCTAATTTGACT 2877  
Db 2761 ATAATGCAATCCTGCTGGGCTTTTGAACCTCAAGGAAACGCCATCTTCCCTAATTTGACT 2820



```
Db 1081 ATTGACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAATCAGA 1140
Qy 1198 TGTACGTGGACCTTCTCTCGAAATCATTTCTCTGAGCAAAAGGCTCTTGATAACGGA 1257
Db 1141 TGTACGTGGACCTTCTCTCGAAATCATTTCTCTGAGCAAAAGGCTCTTGATAACGGA 1200
Qy 1258 TACAGCATATCAAAAGTTTGTCAATCATAAAGCAGCCAGGAGATATATATTCATGCA 1317
Db 1201 TACAGCATATCAAAAGTTTGTCAATCATAAAGCAGCCAGGAGATATATATTCATGCA 1260
Qy 1318 GAAATGATGATGCCCAATTTACCAAAATGTTACAGCTGAGTGAATATAAGAAAGAAACCTCAA 1377
Db 1261 GAAATGATGATGCCCAATTTACCAAAATGTTACAGCTGAGTGAATATAAGAAAGAAACCTCAA 1320
Qy 1378 GTGCTCGCAGAAAGCATCGGCAAGTCAGCGCTCTGTTTCTCGATGGATACCCATTACCA 1437
Db 1321 GTCTCGCAGAAAGCTTGGCAAGTCAGCGCTCTGTTTCTCGATGGATACCCATTACCA 1380
Qy 1438 TCTTGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACAGAA 1497
Db 1381 TCTTGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACAGAA 1440
Qy 1498 GGAGTCTGGAAATAGAAAGGCTAACAGAAAGTGTGTGGACAGTGGGTCTCGAGCAGTACT 1557
Db 1441 GGAGTCTGGAAATAGAAAGGCTAACAGAAAGTGTGTGGACAGTGGGTCTCGAGCAGTACT 1500
Qy 1558 CTAACATGATGAGTGAAGCCATAAAGGTTCTCTGCTCAAGTCTGTGCATACAAATCCCTT 1617
Db 1501 CTAACATGATGAGTGAAGCCATAAAGGTTCTCTGCTCAAGTCTGTGCATACAAATCCCTT 1560
Qy 1618 GGCACATCTTGTGAGACGATCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGAC 1677
Db 1561 GGCACATCTTGTGAGACGATCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGAC 1620
Qy 1678 AACATCTATTTCTATGCAACAATTTGGTGTGTGTCCTCTCTTCATTTGTCGTTTTAACCCGTG 1737
Db 1621 AACATCTATTTCTATGCAACAATTTGGTGTGTGTCCTCTCTTCATTTGTCGTTTTAACCCGTG 1680
Qy 1738 CTAATTTGTACACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG 1797
Db 1681 CTAATTTGTACACAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG 1740
Qy 1798 GTGACCCGGCTCTCAGATAATGAGTACTTCTACCTTGATTTTCAGAGAATATGAATATGAT 1857
Db 1741 GTGACCCGGCTCTCAGATAATGAGTACTTCTACCTTGATTTTCAGAGAATATGAATATGAT 1800
Qy 1858 CTCAAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAGGTTAGGATCAGTGGCT 1917
Db 1801 GTCAAAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAGGTTAGGATCAGTGGCT 1860
Qy 1918 TTTGGAAAAGTGTGATGACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG 1977
Db 1861 TTTGGAAAAGTGTGATGACGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG 1920
Qy 1978 GTTGGCGCTCAAAATGCTGAAAGAAAACAGACAGCTCTGAAAGAGAGGCACTCATGTCA 2037
Db 1921 GTTACCGCTCAAAATGCTGAAAGAAAACAGACAGCTCTGAAAGAGAGGCACTCATGTCA 1980
Qy 2038 GAACTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATATTTGAACCTGCTGGGGGCG 2097
Db 1981 GAACTCAAGATGATGACCCAGCTGGGAAGCCACAGAAATATTTGAACCTGCTGGGGGCG 2040
Qy 2098 TGCACACTGTGACGACCAATTTACTTGATTTTGAATACTGTTGCTATGGTGTCTCTC 2157
Db 2041 TGCACACTGTGACGACCAATTTACTTGATTTTGAATACTGTTGCTATGGTGTCTCTC 2100
Qy 2158 AACTATCTAAGAGTAAAGAGAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAA 2217
Db 2101 AACTATCTAAGAGTAAAGAGAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGGAA 2160
Qy 2218 CACAATTTCAAGTTTAAACCCACTTTTCCAAATCACATCCAAATTCACGATGCCCTGGTTCA 2277
```

```
Db 2161 CACAATTTCAAGTTTAAACCCACTTTTCCAAATCACATCCAAATTCACGATGCTGCTGTCA 2220
Qy 2278 AGAGAAGTTCAGATACACCCGACCTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTT 2337
Db 2221 AGAGAAGTTCAGATACACCCGACCTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTT 2280
Qy 2338 CACTCTGAAGATGAAATTTGAATATGAAAACCAAAAAGGCTGGAAGAAGAGGAGACTTG 2397
Db 2281 CACTCTGAAGATGAAATTTGAATATGAAAACCAAAAAGGCTGGAAGAAGAGGAGACTTG 2340
Qy 2398 AATGCTCTACATTTGAAGATCTCTTGTGCTTTGGATATCAAGTTGCCAAAGAAATGGAA 2457
Db 2341 AATGCTCTACATTTGAAGATCTCTTGTGCTTTGGATATCAAGTTGCCAAAGAAATGGAA 2400
Qy 2458 TTTCTGGAAATTTAAGTCGTGTTTTCACAGAGACTTGGCCGCGCAGGAACGTGCTTGTCAACC 2517
Db 2401 TTTCTGGAAATTTAAGTCGTGTTTTCACAGAGACTTGGCCGCGCAGGAACGTGCTTGTCAACC 2460
Qy 2518 CACGGAAAGTGGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATCATAGTGTATTC 2577
Db 2461 CACGGAAAGTGGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATCATAGTGTATTC 2520
Qy 2578 AACTATGTTGTCAGGGGCAATGCCCGTCTGCCCTGTAAATGGATGGCCCCGCAAGACCTG 2637
Db 2521 AACTATGTTGTCAGGGGCAATGCCCGTCTGCCCTGTAAATGGATGGCCCCGCAAGACCTG 2580
Qy 2638 TTTGAAAGCATCTACACCATTAAGAGTGTCTGTCATATGGAATATTAATCTGTCGGAA 2697
Db 2581 TTTGAAAGCATCTACACCATTAAGAGTGTCTGTCATATGGAATATTAATCTGTCGGAA 2640
Qy 2698 ATCTTCTCACTTGGTGTGAATCCTTACCTTGCATTTCCGGTTTGGCTTCACTTCTACAAA 2757
Db 2641 ATCTTCTCACTTGGTGTGAATCCTTACCTTGCATTTCCGGTTTGGCTTCACTTCTACAAA 2700
Qy 2758 CTGATTTCAAAATGGATTTTAAATGGATCAGCCATTTTATGCTACAGAAATATACATTT 2817
Db 2701 CTGATTTCAAAATGGATTTTAAATGGATCAGCCATTTTATGCTACAGAAATATACATTT 2760
Qy 2818 ATATGCAATCCTGCTGGGCTTTTACCTCAAGGAAACGGCCATCCTTCCCTAATTTGACT 2877
Db 2761 ATATGCAATCCTGCTGGGCTTTTACCTCAAGGAAACGGCCATCCTTCCCTAATTTGACT 2820
Qy 2878 TCGTTTTTAGGATGTGACGTGGCAGATGCAGAAAGGCGATGTATCAGAAATCTGGATGCG 2937
Db 2821 TCGTTTTTAGGATGTGACGTGGCAGATGCAGAAAGGCGATGTATCAGAAATCTGGATGCG 2880
Qy 2938 CGTGTTCGGAAATGTCTCACACCTTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGAT 2997
Db 2881 CGTGTTCGGAAATGTCTCACACCTTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGAT 2940
Qy 2998 TTGGGGCTACTCTCTCCGAGGCTCAGGTGCAAGATTCTGAGAGAAACAAATTTAGTTTFA 3057
Db 2941 TTGGGGCTACTCTCTCCGAGGCTCAGGTGCAAGATTCTGAGAGAAACAAATTTAGTTTFA 3000
Qy 3058 AGGACTTCATCCCTCCACCTATCCCTAACAGCTGTAGATTACCAAAACAAAGATTAATTT 3117
Db 3001 AGGACTTCATCCCTCCACCTATCCCTAACAGCTGTAGATTACCAAAACAAAGATTAATTT 3060
Qy 3118 CATCACTAAAAGAAAATCTATTATCACTGTGCTTCCAGCAGCTTTTCTCTAGAGCGCG 3177
Db 3061 CATCACTAAAAGAAAATCTATTATCACTGTGCTTCCAGCAGCTTTTCTCTAGAGCGCG 3120
```

## RESULT 13

AAV39041

ID AAV39041 standard; cDNA to mRNA; 2949 BP.

XX AAV39041;

AC AAV39041;

XX 12-Oct-1998 (first entry)

XX Human receptor type protein kinase FLT3 encoding cDNA SEQ ID NO.24.

XX







```
Db 1381 TCTTGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCAACTGCACAGAGAGATCACAGAA 1440
Qy 1498 GGAGCTCTGGATACAAAGGCTAACAGAAAAGTGTGGACAGTGGGTGTCAGAGAGTACT 1557
Db 1441 GGAGCTCTGGATACAAAGGCTAACAGAAAAGTGTGGACAGTGGGTGTCAGAGAGTACT 1500
Qy 1558 CTAACATGAGTGAAGCCATAAAGAGGTTCCTGTGCTCAAGTCTGTGCATACAAATCCCTT 1617
Db 1501 CTAACATGAGTGAAGCCATAAAGAGGTTCCTGTGCTCAAGTCTGTGCATACAAATCCCTT 1560
Qy 1618 GGCACATCTTGTGAGAGAGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGAC 1677
Db 1561 GGCACATCTTGTGAGAGAGATCCCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGAC 1620
Qy 1678 AACATCTCATCTATCAACAATTTGGTGTGTGCTCTCTCTCATCTGCTGTTTAAACCTG 1737
Db 1621 AACATCTCATCTATCAACAATTTGGTGTGTGCTCTCTCTCATCTGCTGTTTAAACCTG 1680
Qy 1738 CTAATTTGTCACAAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTTACAG 1797
Db 1681 CTAATTTGTCACAAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTTACAG 1740
Qy 1798 GTGACCGGCTCCTCAGATAATAGTACTTCTACGTTGATTTACAGAGAATATGAATATGAT 1857
Db 1741 GTGACCGGCTCCTCAGATAATAGTACTTCTACGTTGATTTACAGAGAATATGAATATGAT 1800
Qy 1858 CT-----CAAAATGGAGTGTCCAAAGAGAAAATTTA 1887
Db 1801 CTAATTTGTCACAAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGTTACAG 1860
Qy 1888 GAGTTTGGGAGGTACTAGATCAGGTGCTTTTGGAAAAGTGTGAAAGCCAGCTTAT 1947
Db 1861 GAGTTTGGGAGGTACTAGATCAGGTGCTTTTGGAAAAGTGTGAAAGCCAGCTTAT 1920
Qy 1948 GGAATTAGCAAAAAGGAGTCTCAATCCAGGTTCCCGTCAAAATGCTGAAAGAAAAGCA 2007
Db 1921 GGAATTAGCAAAAAGGAGTCTCAATCCAGGTTCCCGTCAAAATGCTGAAAGAAAAGCA 1980
Qy 2008 GACAGCTCTGAAGAGAGGAGGACTCATGTCAAGTCAAGATGATGACCCAGCTGGGAAGC 2067
Db 1981 GACAGCTCTGAAGAGAGGAGGACTCATGTCAAGTCAAGATGATGACCCAGCTGGGAAGC 2040
Qy 2068 CACGAGATAATTTGTAACCTGCTGGGGCGTGCACACTGTGACACCAATTTACTTGTAT 2127
Db 2041 CACGAGATAATTTGTAACCTGCTGGGGCGTGCACACTGTGACACCAATTTACTTGTAT 2100
Qy 2128 TTTGAATACGTGCTATGCTGATCTTCTCAACTATCTAAGAACTAAAGAGAAAATTT 2187
Db 2101 TTTGAATACGTGCTATGCTGATCTTCTCAACTATCTAAGAACTAAAGAGAAAATTT 2160
Qy 2188 CACAGGACTTGGACAGAGATTTTCAAGAGACACAATTTACGTTTACCCGACTTCCAA 2247
Db 2161 CACAGGACTTGGACAGAGATTTTCAAGAGACACAATTTACGTTTACCCGACTTCCAA 2220
Qy 2248 TCACATCCAAATTCACAGCTGCTGGTTCAAGAGAGTTCAGATACACCCGAGCTCGAT 2307
Db 2221 TCACATCCAAATTCACAGCTGCTGGTTCAAGAGAGTTCAGATACACCCGAGCTCGAT 2280
Qy 2308 CAAATCTCAGGGCTTCATGGGAATTTCACTCTGAAAGTGAATTTGAATATGAAAAC 2367
Db 2281 CAAATCTCAGGGCTTCATGGGAATTTCACTCTGAAAGTGAATTTGAATATGAAAAC 2340
Qy 2368 CAAAAAGGCTGGAAGAGAGGAGGACTTGAATGTGCTTACATTTGAAGATCTTCTTTCG 2427
Db 2341 CAAAAAGGCTGGAAGAGAGGAGGACTTGAATGTGCTTACATTTGAAGATCTTCTTTCG 2400
Qy 2428 TTTGCATATCAAGTTGCCAAGGAGTGAATTTCTGAAATTTAAGTGTGTTTCACAGA 2487
Db 2401 TTTGCATATCAAGTTGCCAAGGAGTGAATTTCTGAAATTTAAGTGTGTTTCACAGA 2460
Qy 2488 GACCTGGCCGCCAGGAGGCTGTTGTCACCCAGGAAAGTGGTGAAGATATGACTTT 2547
```

```
Db 2461 GACCTGGCCGCCAGGAAAGCTGCTTGTCAACCCACGGAAGTGGTGAAGATATGTGACTTT 2520
Qy 2548 GGATTGGCTCAGATATATCATGAGTGAATTCCTCAACTATGTTGTGAGGGGCAATGCCCTCTG 2607
Db 2521 GGATTGGCTCAGATATATCATGAGTGAATTCCTCAACTATGTTGTGAGGGGCAATGCCCTCTG 2580
Qy 2608 CCTGTAAATGAGTGGCCCGGAAAGCCTGTTTGAAGGCATCTACACCAATTAAGAGTGAT 2667
Db 2581 CCTGTAAATGAGTGGCCCGGAAAGCCTGTTTGAAGGCATCTACACCAATTAAGAGTGAT 2640
Qy 2668 GTCTGGTCATATGGAATATTTACTGTGGAAATCTTCTCACCTTGTGTGTAATCCTTACCCT 2727
Db 2641 GTCTGGTCATATGGAATATTTACTGTGGAAATCTTCTCACCTTGTGTGTAATCCTTACCCT 2700
Qy 2728 GGCATTCGGTGTGATGCTTAACCTTCTACAAACTGATTTCAAAATGGATTTAAATGGATCAG 2787
Db 2701 GGCATTCGGTGTGATGCTTAACCTTCTACAAACTGATTTCAAAATGGATTTAAATGGATCAG 2760
Qy 2788 CCATTTTATGCTACAGAAGAAATATACATTTATAATGCAATCTGCTGGGCTTTTGACTCA 2847
Db 2761 CCATTTTATGCTACAGAAGAAATATACATTTATAATGCAATCTGCTGGGCTTTTGACTCA 2820
Qy 2848 AGGAAACGGCCATCTTCCCTAAATTTGACTTCGTTTTTAGGATGTCAGCTGGCAGATGCA 2907
Db 2821 AGGAAACGGCCATCTTCCCTAAATTTGACTTCGTTTTTAGGATGTCAGCTGGCAGATGCA 2880
Qy 2908 GAAGAAGCGATGTATCAGAAATGTGATGGCGGTGTTTCGGAATGCTTCCACACCTACCAA 2967
Db 2881 GAAGAAGCGATGTATCAGAAATGTGATGGCGGTGTTTCGGAATGCTTCCACACCTACCAA 2940
Qy 2968 AACAGGCGA 2976
Db 2941 AACAGGCGA 2949

RESULT 14
AAV39040
ID AAV39040 standard; cDNA to mRNA; 2958 BP.
XX AC AAV39040;
XX DT 12-OCT-1998 (first entry)
XX DE Human receptor type protein kinase FLT3 encoding cDNA SEQ ID NO:23.
XX KW Human; receptor type protein kinase; FLT3; diagnosis; leukaemia;
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 1..2958
XX FT /*tag= a
XX FT /note= "no stop codon given"
XX PN W09817808-A1.
XX PD 30-APR-1998.
XX PF 13-OCT-1997; 97WO-JP03667.
XX PR 18-OCT-1996; 96JP-0297329.
XX PA (TAKI ) TAKARA SHUZO CO LTD.
XX PI Yokota S;
XX DR WPI; 1998-362333/31.
XX DR P-PSDB; AAW63587.
XX PT Nucleic acid sequences encoding receptor type protein kinase -
M2 useful for diagnosing adult acute leukaemia of, e.g. FAB sub-classes
```







Db	2221	ACCTTCCAAATCACATCCAAATTCAGCATGCGTGTTCAGAGAAAGTTTCAGATACACCG	2280
Qy	2299	GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTCACCTCGAAGATCAAAATGAA	2358
Db	2281	GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTCACCTCGAAGATCAAAATGAA	2340
Qy	2359	TATGAAACCAAAAGGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT	2418
Db	2341	TATGAAACCAAAAGGCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT	2400
Qy	2419	CTTCTTTCCTTTGCATATCAAGTTGCCAAAGAAATGGAATTTCTGGAATTTAAAGTCGTGT	2478
Db	2401	CTTCTTTCCTTTGCATATCAAGTTGCCAAAGAAATGGAATTTCTGGAATTTAAAGTCGTGT	2460
Qy	2479	GTTCCACAGAGACCTGGCCGCCAGAGAACTGCTGTCCACCCGGAAGAGAGAGAGAGAGAT	2538
Db	2461	GTTCCACAGAGACCTGGCCGCCAGAGAACTGCTGTCCACCCGGAAGAGAGAGAGAGAT	2520
Qy	2539	TGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTCACAACTATGTTGTAGGGGCAAT	2598
Db	2521	TGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTCACAACTATGTTGTAGGGGCAAT	2580
Qy	2599	GCCCGTGTGCTGTAATAATGGATGGCCGCCGAAAGCCCTGTTTGAAGGCATCTACACCAT	2658
Db	2581	GCCCGTGTGCTGTAATAATGGATGGCCGCCGAAAGCCCTGTTTGAAGGCATCTACACCAT	2640
Qy	2659	AAGAGTGATGCTGGTCATATGGATATTAATGAGTGGGAAATCTTCTCACTTGGTGTGAAT	2718
Db	2641	AAGAGTGATGCTGGTCATATGGATATTAATGAGTGGGAAATCTTCTCACTTGGTGTGAAT	2700
Qy	2719	CTTTACCTGGCATTCCGGTTGATGCTAACTTCTACAACTGATTCAAAATGGATTAAA	2778
Db	2701	CTTTACCTGGCATTCCGGTTGATGCTAACTTCTACAACTGATTCAAAATGGATTAAA	2760
Qy	2779	ATGGATCAGCCATTTTATGCTACAGAGAAATATACATTAATGCAATCCTGCTGGGCT	2838
Db	2761	ATGGATCAGCCATTTTATGCTACAGAGAAATATACATTAATGCAATCCTGCTGGGCT	2820
Qy	2839	TTTGACTCAGGAAACGGCCATCCTTCCCTAATTTGACTTCGTTTTAGGATGTCAGCTG	2898
Db	2821	TTTGACTCAGGAAACGGCCATCCTTCCCTAATTTGACTTCGTTTTAGGATGTCAGCTG	2880
Qy	2899	GCAGATGCAGAGAGAGAGAGATGATCAGAAATGATGGATGGCCGTTTTCGGAATGTCCTCAC	2958
Db	2881	GCAGATGCAGAGAGAGAGAGATGATCAGAAATGATGGATGGCCGTTTTCGGAATGTCCTCAC	2940
Qy	2959	ACCTACCAAAACAGGCGA	2976
Db	2941	ACCTACCAAAACAGGCGA	2958

Search completed: May 25, 2003, 12:15:03  
Job time : 710.832 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 04:23:58 ; Search time 146.504 Seconds  
(without alignments)  
7328.626 Million cell updates/sec

Title: US-09-919-408-3  
Perfect score: 3501  
Sequence: 1 CGAGCGCATCCGAGGCT.....AAAAAAAAAAAAAAAAAAAA 3501

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2.6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2.6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2.6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2.6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2.6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2.6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3501	100.0	3501	1	US-07-977-451-3
2	3501	100.0	3501	1	US-08-252-517-3
3	3501	100.0	3501	1	US-07-906-397A-3
4	3501	100.0	3501	1	US-08-601-891-3
5	3501	100.0	3501	2	US-09-021-324-3
6	3501	100.0	3501	5	PCT-US92-05401-3
7	3501	100.0	3501	5	PCT-US92-09893-3
8	3475	99.3	3475	1	US-08-222-299-3
9	3475	99.3	3475	2	US-08-434-878-3
10	3475	99.3	3475	5	PCT-US95-03718-3
11	3474.4	99.2	3476	1	US-08-183-211-1
12	3474.4	99.2	3476	5	PCT-US95-00176A-1
13	3083.2	88.1	3120	1	US-08-222-616-22
14	3083.2	88.1	3120	4	US-08-446-648-22
15	3083.2	88.1	3120	5	PCT-US95-04228-22
16	2274.6	65.0	3521	1	US-08-222-299-1
17	2274.6	65.0	3521	2	US-08-434-878-1
18	2274.6	65.0	3521	5	PCT-US95-03718-1
19	2227.2	63.6	3453	1	US-07-813-593-1
20	2227.2	63.6	3453	1	US-07-977-451-1
21	2227.2	63.6	3453	1	US-07-946-507-1
22	2227.2	63.6	3453	1	US-08-252-517-1
23	2227.2	63.6	3453	1	US-07-906-397A-1
24	2227.2	63.6	3453	1	US-08-601-891-1
25	2227.2	63.6	3453	2	US-09-021-324-1
26	2227.2	63.6	3453	5	PCT-US92-02750-1
27	2227.2	63.6	3453	5	PCT-US92-05401-1

28	2227.2	63.6	3453	5	PCT-US92-09893-1	Sequence 1, Appli
29	624.4	17.8	1894	4	US-07-912-122-3	Sequence 3, Appli
30	624.4	17.8	1894	4	PCT-US93-08404-3	Sequence 3, Appli
31	332	9.5	332	5	PCT-US92-02750-3	Sequence 3, Appli
32	313.2	8.9	5084	1	US-08-306-691B-21	Sequence 21, Appli
33	313.2	8.9	5084	5	PCT-US93-06251-25	Sequence 25, Appli
34	302.6	8.6	4054	1	US-08-180-195-35	Sequence 35, Appli
35	302.6	8.6	4054	1	US-08-477-329-35	Sequence 35, Appli
36	302.6	8.6	4054	2	US-08-475-458-35	Sequence 35, Appli
37	302.6	8.6	4054	3	US-08-980-400-35	Sequence 35, Appli
38	302.6	8.6	4054	4	US-09-583-459A-35	Sequence 35, Appli
39	302.6	8.6	4054	4	US-09-583-210-35	Sequence 35, Appli
40	302.6	8.6	4054	4	US-09-583-449A-35	Sequence 35, Appli
41	302.6	8.6	4054	4	US-09-435-059-35	Sequence 35, Appli
42	302.6	8.6	4100	1	US-08-168-917-3	Sequence 3, Appli
43	302.6	8.6	4100	2	US-08-460-510-3	Sequence 3, Appli
44	302.6	8.6	4100	2	US-08-460-490-3	Sequence 3, Appli
45	302.6	8.6	4100	5	PCT-US92-00730-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1  
US-07-977-451-3  
; Sequence 3, Application US/07977451  
; Patent No. 5270458  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ImClone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/977,451  
; FILING DATE: 19921119  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US UNASSIGNED  
; FILING DATE: 12-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; PRIOR APPLICATION DATA:

```

/ APPLICATION NUMBER: US 07/679,666
/ FILING DATE: 02-APR-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Feit, Irving N.
/ REGISTRATION NUMBER: 28,601
/ REFERENCE/DOCKET NUMBER: LEM-3-7-P
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-645-1405
/ TELEFAX: 212-645-2054
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3501 base pairs
/ TYPE: NUCLEIC ACID
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: N-terminal
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 58..3039
/ FEATURE:
/ NAME/KEY: mat_peptide
/ LOCATION: 139..3036
/ FEATURE:
/ NAME/KEY: sig_peptide
/ LOCATION: 58..138
/ US-07-977-451-3

```

Query Match	100.0%	Score 3501	DB 1	Length 3501
Best Local Similarity	100.0%	Pred. No. 0		
Matches 3501	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	CGAGCGGCATCCAGGCGTGGCGCGCCCTGGGGACCCCGGGCTCCGAGGCCATG	60	
DB	1	CGAGCGGCATCCAGGCGTGGCGCGCCCTGGGGACCCCGGGCTCCGAGGCCATG	60	
QY	61	CCGCGGTTGGCGCGCGACGCGGGGACCGTGCGGTGTTTTCTGCAATGATA	120	
DB	61	CCGCGGTTGGCGCGCGACGCGGGGACCGTGCGGTGTTTTCTGCAATGATA	120	
QY	121	TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCATAG	180	
DB	121	TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTAAATCAATCATAG	180	
QY	181	AACAAATGATTCATCAGTGGGGAAGTCATCATATCCATGGTATCAGAATCCCGGAA	240	
DB	181	AACAAATGATTCATCAGTGGGGAAGTCATCATATCCATGGTATCAGAATCCCGGAA	240	
QY	241	GACCTCGGGTGTGGGTTGAGACCCAGAGCTCAGGGACAGTGTACGAGAGCTGCCGCTGTG	300	
DB	241	GACCTCGGGTGTGGGTTGAGACCCAGAGCTCAGGGACAGTGTACGAGAGCTGCCGCTGTG	300	
QY	301	GAAGTGAGTGTATCGCTTCATCATCACTGCAAGTCTGGTGCATGCCCCAGGGGAACATT	360	
DB	301	GAAGTGAGTGTATCGCTTCATCATCACTGCAAGTCTGGTGCATGCCCCAGGGGAACATT	360	
QY	361	TCCTGCTCTGGGCTTTAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGATTTACAA	420	
DB	361	TCCTGCTCTGGGCTTTAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGATTTACAA	420	
QY	421	AACAGGAGGTGTTTCCATGGTCAATTTTGAANAATGACAGAAACCCAGCTGGAGNATAC	480	
DB	421	AACAGGAGGTGTTTCCATGGTCAATTTTGAANAATGACAGAAACCCAGCTGGAGNATAC	480	
QY	481	CTACTTTTTATTTCAGAGTGAAGCTACCAATTACAAATATCTTTACAGTGAAGTAAAGA	540	
DB	481	CTACTTTTTATTTCAGAGTGAAGCTACCAATTACAAATATCTTTACAGTGAAGTAAAGA	540	
QY	541	AATACCCCTGCTTTACACATTTAAGAGACCTTACTTTAGAAAAATGAAAAACCCAGGACGCC	600	
DB	541	AATACCCCTGCTTTACACATTTAAGAGACCTTACTTTAGAAAAATGAAAAACCCAGGACGCC	600	

QY	601	CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGGAAATGGGTGCTTTGGCAITCA	660
Db	601		
QY	661	CAGGGGAAAGCTGTTAAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGAAAAGTGC	720
Db	661		
QY	721	CATGAATATTTCGGACGCACATAAAGTCCTGCCAGAAATGNACTGGCGCAGGAATGC	780
Db	721		
QY	781	ACCAAGGCTGTTCAAAATAGATCTAAATCAAACTCCTCAGACACACATTCGCACAATATTT	840
Db	781		
QY	841	CTTAAAGTAGGGGAACCCCTTATGGATAAGTGCAAAAGCTGTTCTATGTGAACCATGGATTC	900
Db	841		
QY	901	GGGCTCACCTGGGAATTAGAAAAACAAGCACTCGAGGAGGCAACTTCTTGAGATGAGT	960
Db	901		
QY	961	ACCTATTCAACAACAGAACTATGATACGGATTCTGTTTGGCTTTGTATCATCAGTGGCA	1020
Db	961		
QY	1021	AGAACGACACCCGGATACTACATGTTCTCTTCAAAAGCATCCCAAGTCAATCAGCTTTG	1080
Db	1021		
QY	1081	GTTTACCATTGTAAGAAAGGATTTATAAATGCTACCAATTCAGTGAAGATTTATGAAAT	1140
Db	1081		
QY	1141	GACCAATATGAAGAGTTTGTGTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT	1200
Db	1141		
QY	1201	ACGTGGACCTTCTCTCGAAAATCATTTCCCTTGTGAGCAAAAAGGTCTTGATAACGGATAC	1260
Db	1201		
QY	1261	AGCATATCCAAAGTTTTCGAATATACACACCGCAGGAGAAATATATATTCCTCATGAGAA	1320
Db	1261		
QY	1321	AATGATGATGCCAAATTTACCAAAATGTTACGCTGAATATAGAGGAAACCTCAAGTG	1380
Db	1321		
QY	1381	CTCGCAGAAGCATCGCAAGTCAGGCGTCTGTTTCTCGGATGGATACCCATTACCATCT	1440
Db	1381		
QY	1441	TGGACCTGGAGAAGTGTTCAGACAAGTCTCCCAACTGCACAGNAGAGATCAGAGGA	1500
Db	1441		
QY	1501	GTCTGGAATAGAAAGGCTTAACAGAAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA	1560
Db	1501		
QY	1561	RACATGAGTGAAGCCATAAAAAGGTTTCCTGCTCAAGTGTGTGCATACAAATTCCTTTGGC	1620
Db	1561		
QY	1621	ACATCTGTGAGACGATCCCTTTTAACTCTCCAGCCCTTCCCTTTTCAATCCACAGACAAC	1680
Db	1621		

- 9 -

Qy	1681	ATCTCATTCATGCAACAACAAATTGGTGTGTTTGTCCTCCTCTCAATGTCGTGTTTTAAACCCCTGCCTA	1740
Db	1681	ATCTCATTCATGCAACAACAAATTGGTGTGTTTGTCCTCCTCTCAATGTCGTGTTTTAAACCCCTGCCTA	1740
Qy	1741	ATTTGTGCACAAGTACAAAAGCAAAATTAGGTATGAAGCCAGCTACAGATGGTACAGGTG	1800
Db	1741	ATTTGTGCACAAGTACAAAAGCAAAATTAGGTATGAAGCCAGCTACAGATGGTACAGGTG	1800
Qy	1801	ACCGGCTCCTCAGATAATGAGTAGTACTTCTACGTTGATTTAGAGAATATGAATATGATCTC	1860
Db	1801	ACCGGCTCCTCAGATAATGAGTAGTACTTCTACGTTGATTTAGAGAATATGAATATGATCTC	1860
Qy	1861	AATATGGAGTTTCCAAGAGAAAAATTAGAGTTTGGGAAGTACTAGATCAGTGCTTTT	1920
Db	1861	AATATGGAGTTTCCAAGAGAAAAATTAGAGTTTGGGAAGTACTAGATCAGTGCTTTT	1920
Qy	1921	GGAAGAAGTATGAACCAACAGACTTATGGAAATTAGCAAACAGAGTCTCAATCCAGGTT	1980
Db	1921	GGAAGAAGTATGAACCAACAGACTTATGGAAATTAGCAAACAGAGTCTCAATCCAGGTT	1980
Qy	1981	GCCGCTCAAAATGCTGAAGAAAAAGCAGACAGCTCTCAAGAGAGGCACTCATGTGCAGAA	2040
Db	1981	GCCGCTCAAAATGCTGAAGAAAAAGCAGACAGCTCTCAAGAGAGGCACTCATGTGCAGAA	2040
Qy	2041	CTCAAGATGATACCAGCTGGGAAGCCACGAGAAATATGTGAACCTGCTGGGGGGCTGC	2100
Db	2041	CTCAAGATGATACCAGCTGGGAAGCCACGAGAAATATGTGAACCTGCTGGGGGGCTGC	2100
Qy	2101	ACACTGTCAGGCCAATTACTTTGATTTTGAATACTGTGTCTATGGTGATCTTCTCAAC	2160
Db	2101	ACACTGTCAGGCCAATTACTTTGATTTTGAATACTGTGTCTATGGTGATCTTCTCAAC	2160
Qy	2161	TATCTAAGAAGTAAAAAGAGAAAAAATTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC	2220
Db	2161	TATCTAAGAAGTAAAAAGAGAAAAAATTTCACAGGACTTGGACAGAGATTTTCAAGGAACAC	2220
Qy	2221	AATTTACAGTTTACCCCACTTTCCAAATCACATCCAAATTTCCASGACTGCTGGTCTCAAGA	2280
Db	2221	AATTTACAGTTTACCCCACTTTCCAAATCACATCCAAATTTCCASGACTGCTGGTCTCAAGA	2280
Qy	2281	GAAGTTTCAGATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAAC	2340
Db	2281	GAAGTTTCAGATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAAC	2340
Qy	2341	TCCTGAAGATGAAATTTGAATATGAAAAACAAAAAGGCTGGAAGAAGAGGAGGACTTTGAAT	2400
Db	2341	TCCTGAAGATGAAATTTGAATATGAAAAACAAAAAGGCTGGAAGAAGAGGAGGACTTTGAAT	2400
Qy	2401	GTGCTTACATTTGAAGATCTTCTTTGGCTTTGCATATCAAGTTGCCAAGGAATGGAAATTT	2460
Db	2401	GTGCTTACATTTGAAGATCTTCTTTGGCTTTGCATATCAAGTTGCCAAGGAATGGAAATTT	2460
Qy	2461	CTGGAATTTAAGTCGTGTTCACAGAGACCTGGCCCCAGGAACGCTGCTGTGTCACCCCAC	2520
Db	2461	CTGGAATTTAAGTCGTGTTCACAGAGACCTGGCCCCAGGAACGCTGCTGTGTCACCCCAC	2520
Qy	2521	GGGAAGTGGTGAAGATATCTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTTCCAAC	2580
Db	2521	GGGAAGTGGTGAAGATATCTGACTTTGGATTTGGCTCGAGATATCATGAGTGATTTCCAAC	2580
Qy	2581	TATGTTGTTCAGGGGCAATGCCCGCTGCCCTGTAAAATGGATGGCCCCGGAAGGCCCTGTTT	2640
Db	2581	TATGTTGTTCAGGGGCAATGCCCGCTGCCCTGTAAAATGGATGGCCCCGGAAGGCCCTGTTT	2640
Qy	2641	GAAGGCATCTACACCATTAGAGTGATGTCGTGCATATGGAAATTTACTTGTGGGAATC	2700
Db	2641	GAAGGCATCTACACCATTAGAGTGATGTCGTGCATATGGAAATTTACTTGTGGGAATC	2700
Qy	2701	TTCTCACTTGGTGTGAATCCTTACCCTGGCAATTCGGGTTGATGCTAACCTTCTACAAACTG	2760
Db	2701	TTCTCACTTGGTGTGAATCCTTACCCTGGCAATTCGGGTTGATGCTAACCTTCTACAAACTG	2760
Qy	2761	ATTCAAAATGGATTTAAAAATGGATCAGCCCAATTTTATGCTACAGAAGAAATATACATTAATA	2820

Db	2761		ATTCAAATGGATTTAAATGGATGACCCATTTATGTCTCAGAGAAATATACATTTATA	2820
Qy	2821		ATGCAATCCCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCCTTCCCTAAATTTGACTTCG	2880
Db	2821		ATGCAATCCCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCCTTCCCTAAATTTGACTTCG	2880
Qy	2881		TTTTTAGTAGTGTCAAGCTGGCAGATGCAAGAAAGCGATGTATCAGAAATGTGGATGGCCGT	2940
Db	2881		TTTTTAGTAGTGTCAAGCTGGCAGATGCAAGAAAGCGATGTATCAGAAATGTGGATGGCCGT	2940
Qy	2941		GTTTCGGAATGCTCTACACCTACCAAAACAGGGGACCTTTTCAGCAGAGAGATGGATTTG	3000
Db	2941		GTTTCGGAATGCTCTACACCTACCAAAACAGGGGACCTTTTCAGCAGAGAGATGGATTTG	3000
Qy	3001		GGGCTACTCTCTCGCAGGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG	3060
Db	3001		GGGCTACTCTCTCGCAGGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG	3060
Qy	3061		ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTTTCAT	3120
Db	3061		ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTTTCAT	3120
Qy	3121		CACATAAAGAAAAATCTATTATCAACTGCTGCTTCCACCAGACTTTTCTCTAGAAGCCGTCT	3180
Db	3121		CACATAAAGAAAAATCTATTATCAACTGCTGCTTCCACCAGACTTTTCTCTAGAAGCCGTCT	3180
Qy	3181		CGGTTTACTCTGTGTTTCAAAGGACCTTTGTAAAATCAAATCATCTGTCAACAAGGCAG	3240
Db	3181		CGGTTTACTCTGTGTTTCAAAGGACCTTTGTAAAATCAAATCATCTGTCAACAAGGCAG	3240
Qy	3241		GAGGAGCTGATATGAACCTTTATGGAGCATTTGATCTGCATCCAAAGGCCCTCTCAGGCCG	3300
Db	3241		GAGGAGCTGATATGAACCTTTATGGAGCATTTGATCTGCATCCAAAGGCCCTCTCAGGCCG	3300
Qy	3301		GCTTGAGTGAATCTGTACCTGGAAGTACAGTATATCTTGTAAATACATAAAACAAGGC	3360
Db	3301		GCTTGAGTGAATCTGTACCTGGAAGTACAGTATATCTTGTAAATACATAAAACAAGGC	3360
Qy	3361		ATTTTGCTAAGGAGAAGCTAATATGATTTTTTAAAGTCTATGTTTTAAAAATAATGTAAA	3420
Db	3361		ATTTTGCTAAGGAGAAGCTAATATGATTTTTTAAAGTCTATGTTTTAAAAATAATGTAAA	3420
Qy	3421		TTTTTTCAGCTATTTAGTGATATATTTTATGGGTGGGAATAAAAATTTCTACTACAGAAAA	3480
Db	3421		TTTTTTCAGCTATTTAGTGATATATTTTATGGGTGGGAATAAAAATTTCTACTACAGAAAA	3480
Qy	3481		AAAAAAAAAAAAAAAAAAAAA 3501	
Db	3481		AAAAAAAAAAAAAAAAAAAAA 3501	

```

RESULT 2
US-08-252-517-3
; Sequence 3, Application US/08252517
; Patent No. 5548065
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```



CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/252,517  
FILING DATE: 31-OCT-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/977,451  
FILING DATE: 19-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3501 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 58..3039  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 139..3036  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 58..138  
US-08-252-517-3

Query Match 100.0%; Score 3501; DB 1; Length 3501;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CGAGCGGCATCCGAGGGCTGGCGCGGCCCTGGGGACCCCGGCTCCGGAGGCCATG 60  
Db 1 CGAGCGGCATCCGAGGGCTGGCGCGGCCCTGGGGACCCCGGCTCCGGAGGCCATG 60  
Qy 61 CGCGGCTTGGCGCGGACCGGCGACCGTGGCGCTGCTGCTGTTTTTCTGCAATGATA 120  
Db 61 CGCGGCTTGGCGCGGACCGGCGACCGTGGCGCTGCTGCTGTTTTTCTGCAATGATA 120  
Qy 121 TTGGGACTATTACAAATCAAGATCGCTGTCATCAAGTGTCTTTAATCAATCAATAG 180  
Db 121 TTGGGACTATTACAAATCAAGATCGCTGTCATCAAGTGTCTTTAATCAATCAATAG 180

121 TTGGGACTATTACAAATCAAGATCGCTGTCATCAAGTGTCTTTAATCAATCAATAG 180  
181 AACATGATTCATCAGTGGGGAAGTCATCATATCCATGTCATCAGATCCCGGAA 240  
181 AACATGATTCATCAGTGGGGAAGTCATCATATCCATGTCATCAGATCCCGGAA 240  
241 GACCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGACAGTGTAGGAACTGCCGCTG 300  
241 GACCTCGGGTGTGCGTTGAGACCCAGAGCTCAGGACAGTGTAGGAACTGCCGCTG 300  
301 GAAGTGGATGATCTGCTTCCATCATACATGCAAGTGTGCTGATGCCCGGAGACAT 360  
301 GAAGTGGATGATCTGCTTCCATCATACATGCAAGTGTGCTGATGCCCGGAGACAT 360  
361 TCCTGCTCTGGTCTTTAAGCAGACGCTCCCTGAAATGCGACGCCACATTTGATTTACAA 420  
361 TCCTGCTCTGGTCTTTAAGCAGACGCTCCCTGAAATGCGACGCCACATTTGATTTACAA 420  
421 AACAGAGGAGTTGTTTCCATGCTCATTTTGAATAACAGAGAACCCCAAGCTGGAGAAATAC 480  
421 AACAGAGGAGTTGTTTCCATGCTCATTTTGAATAACAGAGAACCCCAAGCTGGAGAAATAC 480  
481 CTACTTTTATTCAGAGTGAAGCTACCAATACACAATATTTTACAGTGTAGTATAAGA 540  
481 CTACTTTTATTCAGAGTGAAGCTACCAATACACAATATTTTACAGTGTAGTATAAGA 540  
541 AATACCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGAAAACACGAGCGCC 600  
541 AATACCTGCTTTACACATTAAGAGACCTTACTTTAGAAAATGAAAACACGAGCGCC 600  
601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGAATGGTGTCTTCGATTCA 660  
601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCGGATCGTGAATGGTGTCTTCGATTCA 660  
661 CAGGGGAAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTGCCT 720  
661 CAGGGGAAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGGAAAGTGCCT 720  
721 CATGAATTTTGGGACGACATAGGTGCTGCCAGAAATGAACCTGGCAGGAGATGC 780  
721 CATGAATTTTGGGACGACATAGGTGCTGCCAGAAATGAACCTGGCAGGAGATGC 780  
781 ACCAGGCTGTTCCACATAGATCTAAATCAAACTCCCTCAGACCACTATGTCACAAATATTT 840  
781 ACCAGGCTGTTCCACATAGATCTAAATCAAACTCCCTCAGACCACTATGTCACAAATATTT 840  
841 CTTAAAGTAGGGGAACCCCTTATGGATGAAGTGCAGAACTGTTTCATGTGAACCATGGATTC 900  
841 CTTAAAGTAGGGGAACCCCTTATGGATGAAGTGCAGAACTGTTTCATGTGAACCATGGATTC 900  
901 GGGCTCACCTGGGAATTAGAAAACAAAGCACATCGAGGAGGGCACTACTTTGAGATCAGT 960  
901 GGGCTCACCTGGGAATTAGAAAACAAAGCACATCGAGGAGGGCACTACTTTGAGATCAGT 960  
961 ACCTATTCAACAAACAGAACTATGATACGATGCTGTTGCTTTGTTATCATCAGTGGCA 1020  
961 ACCTATTCAACAAACAGAACTATGATACGATGCTGTTGCTTTGTTATCATCAGTGGCA 1020  
1021 AGAAGCAGACCCGGATACACTTGTCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080  
1021 AGAAGCAGACCCGGATACACTTGTCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080  
1081 GTTACCATCGTAGGAAAGGATTTAATGCTACCAATTCAGTGAAGTATGAAT 1140  
1081 GTTACCATCGTAGGAAAGGATTTAATGCTACCAATTCAGTGAAGTATGAAT 1140  
1141 GACCAATACAGAGTTTCTGTTTCTGTCAGGTTTAAAGCCCTACCCACAAATCAGATGT 1200  
1141 GACCAATACAGAGTTTCTGTTTCTGTCAGGTTTAAAGCCCTACCCACAAATCAGATGT 1200  
1201 ACCTGGACCTTCTCTGAAAATCATTTCTTGTGAGCAAAAGGCTCTTGAACGGATAC 1260  
1201 ACCTGGACCTTCTCTGAAAATCATTTCTTGTGAGCAAAAGGCTCTTGAACGGATAC 1260

1261 AGCATATCCAAGTTTTCGAATCATAGCACACAGCCAGGAGAATATATATTCCTTCAGAA 1320  
1261 AGCATATCCAAGTTTTCGAATCATAGCACACAGCCAGGAGAATATATATTCCTTCAGAA 1320  
1321 AATGATGATGCCCAATTTACCAAAATGTTTCCAGCTGAAATATAGAGAGAAACCTTCAAGTG 1380  
1321 AATGATGATGCCCAATTTACCAAAATGTTTCCAGCTGAAATATAGAGAGAAACCTTCAAGTG 1380  
1381 CTCGCAGAGCATCGCAAGTCAGGCGTCTCTGTTCTCGGATGGATACCCATTAACCATCT 1440  
1381 CTCGCAGAGCATCGCAAGTCAGGCGTCTCTGTTCTCGGATGGATACCCATTAACCATCT 1440  
1441 TGGACCTGGAAGAGTGTTCAGACAAGTCFCCAACTGTCACAGAAGAGATCACAGAAGGA 1500  
1441 TGGACCTGGAAGAGTGTTCAGACAAGTCFCCAACTGTCACAGAAGAGATCACAGAAGGA 1500  
1501 GTCTGGAATAGAAAGGCTAACAGAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
1501 GTCTGGAATAGAAAGGCTAACAGAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACAAATTCCTTTGGC 1620  
1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACAAATTCCTTTGGC 1620  
1621 ACATCTGTGAGAGATCCCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGACAAC 1680  
1621 ACATCTGTGAGAGATCCCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGACAAC 1680  
1681 ATCTCATCTTATCAACAATTTGGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTA 1740  
1681 ATCTCATCTTATCAACAATTTGGTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTA 1740  
1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
1801 ACCGGCTCCTCAGATATGATGACTCTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860  
1801 ACCGGCTCCTCAGATATGATGACTCTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860  
1861 AAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAGGTACTPAGGATCAGGTGCTTTT 1920  
1861 AAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAGGTACTPAGGATCAGGTGCTTTT 1920  
1921 GGAAGTATGATGACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
1921 GGAAGTATGATGACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACCTCATCTCAGAA 2040  
1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACCTCATCTCAGAA 2040  
2041 CTCGAAGTATGATGCCAGCTGGGAAGCCACGAGAATATTTGGAACCTGCTGGGGCGCTGC 2100  
2041 CTCGAAGTATGATGCCAGCTGGGAAGCCACGAGAATATTTGGAACCTGCTGGGGCGCTGC 2100  
2101 ACATGTCAGACCAATTTACTTGATTTTGAATPACTGTTGCTATGTTGATCTCTCAAC 2160  
2101 ACATGTCAGACCAATTTACTTGATTTTGAATPACTGTTGCTATGTTGATCTCTCAAC 2160  
2161 TATCTAGAGATGAAAGAGAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGAAACAC 2220  
2161 TATCTAGAGATGAAAGAGAAAATTTTCACAGGACTTGGACAGAGATTTTCAAGAAACAC 2220  
2221 AATTTCAAGTTTATACCCCACTTTCCAATCATCATCCAAATTCACGATGCCCTGGTTCAAGA 2280  
2221 AATTTCAAGTTTATACCCCACTTTCCAATCATCATCCAAATTCACGATGCCCTGGTTCAAGA 2280  
2281 GAAGTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAAC 2340  
2281 GAAGTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAAC 2340

2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAAGGCTGGGAAGAGGAGGACTTGAAT 2400  
2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAAGGCTGGGAAGAGGAGGACTTGAAT 2400  
2401 GTGCTTACATTTGAAGATCTCTTTGCTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
2401 GTGCTTACATTTGAAGATCTCTTTGCTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
2461 CTGGAATTTAAGTCGTGTTCACAGAGACCTGGCCGACGAGAAAGCTGTTGTACCCAC 2520  
2461 CTGGAATTTAAGTCGTGTTCACAGAGACCTGGCCGACGAGAAAGCTGTTGTACCCAC 2520  
2521 GGGAAAGTGTGGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTTCCAAC 2580  
2521 GGGAAAGTGTGGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTTCCAAC 2580  
2581 TATGTTGTCAGGGCAATGCCCGTCTGCTGTAAAATGGATGGCCCGGAAAGCCTGTTT 2640  
2581 TATGTTGTCAGGGCAATGCCCGTCTGCTGTAAAATGGATGGCCCGGAAAGCCTGTTT 2640  
2641 GAAGGCATCTACACCAATTAAGAGTATGCTGTGGTCATATGGAATATTTACTGTGGGAATC 2700  
2641 GAAGGCATCTACACCAATTAAGAGTATGCTGTGGTCATATGGAATATTTACTGTGGGAATC 2700  
2701 TTCTCAGCTTGGTGTGAATTCCTTACCTGGCATTTCCGGTTGATGCTTAACCTTCTACAACTG 2760  
2701 TTCTCAGCTTGGTGTGAATTCCTTACCTGGCATTTCCGGTTGATGCTTAACCTTCTACAACTG 2760  
2761 ATTCAAATGATTTTAAATGGATCAGCCATTTTATGTACAGAAGAATATACATATATA 2820  
2761 ATTCAAATGATTTTAAATGGATCAGCCATTTTATGTACAGAAGAATATACATATATA 2820  
2821 ATGCAATCCTCTCGGCTTTTACTCAAGAAACGGCCATCTTCCCTAAATTTGACTTCG 2880  
2821 ATGCAATCCTCTCGGCTTTTACTCAAGAAACGGCCATCTTCCCTAAATTTGACTTCG 2880  
2881 TTTTATGATCTCAGCTGGCAGATGCAGAAAGCGATGTATCAGAATGTGGATGGCGCT 2940  
2881 TTTTATGATCTCAGCTGGCAGATGCAGAAAGCGATGTATCAGAATGTGGATGGCGCT 2940  
2941 GTTTCGGAATGCTCTCACACCTACAAAACAGGGACCTTTTCAGCAGAGAGATGATTTG 3000  
2941 GTTTCGGAATGCTCTCACACCTACAAAACAGGGACCTTTTCAGCAGAGAGATGATTTG 3000  
3001 GGGCTACTCTCTCGCAGGCTCAGGTCGAAGATTCGTAGAGAGAAATTTAGTTTAAAG 3060  
3001 GGGCTACTCTCTCGCAGGCTCAGGTCGAAGATTCGTAGAGAGAAATTTAGTTTAAAG 3060  
3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAGATTAATTTAT 3120  
3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAGATTAATTTAT 3120  
3121 CACTAAAAAGAAAATCTATTATCAACTGCTGCTTACCAGACTTTTCTCTAGAACGCGTCT 3180  
3121 CACTAAAAAGAAAATCTATTATCAACTGCTGCTTACCAGACTTTTCTCTAGAACGCGTCT 3180  
3181 CGGTTTACTCTGTTTTCAAAAGGACTTTTGTAAAATCAATCATCTCTGTCACAGGACG 3240  
3181 CGGTTTACTCTGTTTTCAAAAGGACTTTTGTAAAATCAATCATCTCTGTCACAGGACG 3240  
3241 GAGGAGCTGATTAATGAACCTTTTATCGAGCATTTGATCTCATCCAGGCTTCTCTAGGCCG 3300  
3241 GAGGAGCTGATTAATGAACCTTTTATCGAGCATTTGATCTCATCCAGGCTTCTCTAGGCCG 3300  
3301 GCTTGAAGTAAATGCTGTAACCTGAGTACAGTATATTTTGTAAAATACATAAAACAAAAGC 3360  
3301 GCTTGAAGTAAATGCTGTAACCTGAGTACAGTATATTTTGTAAAATACATAAAACAAAAGC 3360  
3361 ATTTTGTGTAAGAGAGCTAATATGATTTTAACTCTATGTTTAAATAATATATGTA 3420  
3361 ATTTTGTGTAAGAGAGCTAATATGATTTTAACTCTATGTTTAAATAATATATGTA 3420  
3421 TTTTTCAGCTATTTAGTATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAA 3480

```

Db 3421 TTTTTCAGCTATTAGTGATATATTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480
Qy 3481 AAAAAAAAAAAAAAAAAAAAA 3501
Db 3481 AAAAAAAAAAAAAAAAAAAAA 3501

RESULT 3
US-07-906-397A-3
; Sequence 3, Application US/07906397A
; Patent No. 5621090
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Thor R.
; TITLE OF INVENTION: TOPIPORENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSER: IMCLONE SYSTEMS INCORPORATED
; STREET: 180 VARICK STREET
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,397A
; FILING DATE: 19920626
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3501 base pairs
; TYPE: NUCLEIC ACID
; TOPOLOGY: linear
; STRANDEDNESS: single
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 58..3039
US-07-906-397A-3

Query Match 100.0%; Score 3501; DB 1; Length 3501;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGAGCGGCATCCGAGGGCTGGCGCGCCCTGGGGGACCCCGGCTCCGGAGGCCATG 60
Db 1 CGAGCGGCATCCGAGGGCTGGCGCGCCCTGGGGGACCCCGGCTCCGGAGGCCATG 60

```

```

Qy 61 CCGGCGTTGGCGCGCACGGGCGCACCGTCCGCTGCTGTTGTTTTCTGCAATGATA 120
Db 61 CCGGCGTTGGCGCGCACGGGCGCACCGTCCGCTGCTGTTGTTTTCTGCAATGATA 120
Qy 121 TTTGGGACTATTACAAATCAAGATCTGCCCTGTGATCAAGTGTGTTTTTAATCAATCAAG 180
Db 121 TTTGGGACTATTACAAATCAAGATCTGCCCTGTGATCAAGTGTGTTTTTAATCAATCAAG 180
Qy 181 AACAAATGATTATCAGTGGGAAAGTATCATATATCCCATGTGATCAGAAATCCCGGAA 240
Db 181 AACAAATGATTATCAGTGGGAAAGTATCATATATCCCATGTGATCAGAAATCCCGGAA 240
Qy 241 GACCTCGGGTGTGCTTGACACCCAGAGCTCAGGGACAGTGTACGAGCTGCCGCTGTG 300
Db 241 GACCTCGGGTGTGCTTGACACCCAGAGCTCAGGGACAGTGTACGAGCTGCCGCTGTG 300
Qy 301 GAAGTGGATGTATCTGCTTCCATCACACTGCTGCTGATGCTGCCAGGGAACATT 360
Db 301 GAAGTGGATGTATCTGCTTCCATCACACTGCTGCTGATGCTGCCAGGGAACATT 360
Qy 361 TCCTGTCTCTGGGTCTTTAAGCACAGCTCCCTGAATTCGCCAGCCACATTTTGAATACAA 420
Db 361 TCCTGTCTCTGGGTCTTTAAGCACAGCTCCCTGAATTCGCCAGCCACATTTTGAATACAA 420
Qy 421 AACAGAGGAGTTGTTCCATGTTGTTTGAATATGACAGAAACCCAGCTGGAGAATAC 480
Db 421 AACAGAGGAGTTGTTCCATGTTGTTTGAATATGACAGAAACCCAGCTGGAGAATAC 480
Qy 481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTGTACAGTCAAGTATAGA 540
Db 481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTGTACAGTCAAGTATAGA 540
Qy 541 AATACCTGCTTTACACATTAAAGAAGACCTTACTTTAGAAAAATGAAAAACCCAGAGCC 600
Db 541 AATACCTGCTTTACACATTAAAGAAGACCTTACTTTAGAAAAATGAAAAACCCAGAGCC 600
Qy 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCTGGAATGGGTGCTTTGGATGATCA 660
Db 601 CTGGTCTGCATATCTGAGAGCGTTCCAGAGCCGATCTGGAATGGGTGCTTTGGATGATCA 660
Qy 661 CAGGGGAAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTAAAAAGGAGAAAAAGTGCCT 720
Db 661 CAGGGGAAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTAAAAAGGAGAAAAAGTGCCT 720
Qy 721 CATCAATTTATTTGGACGGACATAAGGTGCTGCCAGAAATGAACCTGGGAGGAATGC 780
Db 721 CATCAATTTATTTGGACGGACATAAGGTGCTGCCAGAAATGAACCTGGGAGGAATGC 780
Qy 781 ACCAGGCTGTTCAACAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCACAATTTT 840
Db 781 ACCAGGCTGTTCAACAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCACAATTTT 840
Qy 841 CTTAAAGTAGGGAAACCTTTATGGATAAGGTGCAAAAGCTGTTCAATGTGAACCATGGATT 900
Db 841 CTTAAAGTAGGGAAACCTTTATGGATAAGGTGCAAAAGCTGTTCAATGTGAACCATGGATT 900
Qy 901 GGGCTCACCTGGGAATTTAGAAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT 960
Db 901 GGGCTCACCTGGGAATTTAGAAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT 960
Qy 961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTTGTGTTTGTATCATCATAGTGCA 1020
Db 961 ACCTATTCAACAAACAGAACTATGATACGGATTCTGTTTGTGTTTGTATCATCATAGTGCA 1020
Qy 1021 AGAACGACACCGGATACACTGTTTCCCTTTCAAGCATCCAGTCAATAGCTTTG 1080
Db 1021 AGAACGACACCGGATACACTGTTTCCCTTTCAAGCATCCAGTCAATAGCTTTG 1080
Qy 1081 GTTACCATCTAGGAAAGGATTTTATAAATGCTACCAATTTCAAGTGAAGATTTGAAAT 1140
Db 1081 GTTACCATCTAGGAAAGGATTTTATAAATGCTACCAATTTCAAGTGAAGATTTGAAAT 1140

```

QY 1141 GACCAATATGAGAGTTTGTCTTCTGTGTCAGGTTTAAAGCCTTACCACAAATCAGATGT 1200  
Db 1141 GACCAATATGAGAGTTTGTCTTCTGTGTCAGGTTTAAAGCCTTACCACAAATCAGATGT 1200  
QY 1201 ACCTGGACCTTCTCTCGAAATCATTTCCCTGTGTGAGCAAAAGGTCGTGTATACGGTAC 1260  
Db 1201 ACCTGGACCTTCTCTCGAAATCATTTCCCTGTGTGAGCAAAAGGTCGTGTATACGGTAC 1260  
QY 1261 AGCATATCCAAAGTTTGTCAATCATAGCACAGCAGGAGAAATATATATCCATGTCAGAA 1320  
Db 1261 AGCATATCCAAAGTTTGTCAATCATAGCACAGCAGGAGAAATATATATCCATGTCAGAA 1320  
QY 1321 AATGATGATGCCCAATTTACCAAAATGTTCCACGCTGAATATATAGAAAGAAACCTTCAAGTG 1380  
Db 1321 AATGATGATGCCCAATTTACCAAAATGTTCCACGCTGAATATATAGAAAGAAACCTTCAAGTG 1380  
QY 1381 CTGCGAGAAGCATCGGCAAGTCAGGCGTCTGTTCTCGATGGATACCATTTACCATCT 1440  
Db 1381 CTGCGAGAAGCATCGGCAAGTCAGGCGTCTGTTCTCGATGGATACCATTTACCATCT 1440  
QY 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCAACTGCCAGAGAGATCACAGAAGGA 1500  
Db 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCAACTGCCAGAGAGATCACAGAAGGA 1500  
QY 1501 GTCTGGAATAGAAAGGCTAACAGAAAGTGTGTCAGAGTGGGTGTGCGAGCAGTACTCTA 1560  
Db 1501 GTCTGGAATAGAAAGGCTAACAGAAAGTGTGTCAGAGTGGGTGTGCGAGCAGTACTCTA 1560  
QY 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTGCATACAAATCCCTTGGC 1620  
Db 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTGCATACAAATCCCTTGGC 1620  
QY 1621 ACATCTGTGAGAGGATCCCTTTAACTCTCCAGGCCCTTCCCTTTCATCCCAAGACAAC 1680  
Db 1621 ACATCTGTGAGAGGATCCCTTTAACTCTCCAGGCCCTTCCCTTTCATCCCAAGACAAC 1680  
QY 1681 ATCTCATCTATGCAAAATGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
Db 1681 ATCTCATCTATGCAAAATGTTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
QY 1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Db 1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
QY 1801 ACCGGCTCTCAGATATAGTACTCTACGTTGATTTTCAGAGAAATATGATATGATCTC 1860  
Db 1801 ACCGGCTCTCAGATATAGTACTCTACGTTGATTTTCAGAGAAATATGATATGATCTC 1860  
QY 1861 AATGGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
Db 1861 AATGGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
QY 1921 GGAAGAGTGAAGCAAGCAAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
Db 1921 GGAAGAGTGAAGCAAGCAAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
QY 1981 GCGCTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGGCACTCATCTCAGAA 2040  
Db 1981 GCGCTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGGCACTCATCTCAGAA 2040  
QY 2041 CTCAGATGATGCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGCGGTGC 2100  
Db 2041 CTCAGATGATGCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGCGGTGC 2100  
QY 2101 ACAGTGTGAGGACCAATTTACTGTATTTTGAATGATGTTGTGCTATGATGATCTCTCAAC 2160  
Db 2101 ACAGTGTGAGGACCAATTTACTGTATTTTGAATGATGTTGTGCTATGATGATCTCTCAAC 2160  
QY 2161 TATCTAAGAGTAAAGAGAAAATTTTCAGAGNCTTGGACAGAGATTTTTCAGGACAC 2220  
Db 2161 TATCTAAGAGTAAAGAGAAAATTTTCAGAGNCTTGGACAGAGATTTTTCAGGACAC 2220  
QY 2221 AATTTTCAGTTTACCCTCACTTCCCAATCATCAATCCCAATCCAGCATGCTGGTTCAAGA 2280

Db 2221 AATTTTCAGTTTACCCTCACTTCCCAATCATCAATCCCAATCCAGCATGCTGGTTCAAGA 2280  
QY 2281 GAAGTTGAGATACACCGGACTCGGATCAAAATCTCAGGCTTCATGGGAATTCATTTCAC 2340  
Db 2281 GAAGTTGAGATACACCGGACTCGGATCAAAATCTCAGGCTTCATGGGAATTCATTTCAC 2340  
QY 2341 TCTCAAGATGAAATTTGAATATGAAACCAAAAGGCTGGGAAGAGAGAGGAGACTTGAAT 2400  
Db 2341 TCTCAAGATGAAATTTGAATATGAAACCAAAAGGCTGGGAAGAGAGAGGAGACTTGAAT 2400  
QY 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTGCATATCAAGTTGCCAAAGAAATTT 2460  
Db 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTGCATATCAAGTTGCCAAAGAAATTT 2460  
QY 2461 CTGGAATTTAAGTCTGTGTTTCACAGACACCTGGCCGCGAGGAAGTCTTCTCACCAC 2520  
Db 2461 CTGGAATTTAAGTCTGTGTTTCACAGACACCTGGCCGCGAGGAAGTCTTCTCACCAC 2520  
QY 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGCTTCGAGATATCATGAGTCAATCCAA 2580  
Db 2521 GGGAAAGTGGTGAAGATATGTGACTTTGGCTTCGAGATATCATGAGTCAATCCAA 2580  
QY 2581 TATGTTGTCAGGGCAATGCCCTGTGCTGTAAAATGATGGCCCGGAAAGCCCTGTTT 2640  
Db 2581 TATGTTGTCAGGGCAATGCCCTGTGCTGTAAAATGATGGCCCGGAAAGCCCTGTTT 2640  
QY 2641 GAAGCATCTACACCATTAAGAGTATGATCTGCTGATGGAATATTTACTGTGGGAATC 2700  
Db 2641 GAAGCATCTACACCATTAAGAGTATGATCTGCTGATGGAATATTTACTGTGGGAATC 2700  
QY 2701 TTCTCACTTGTGTGAATCCCTTACCCTGGCATTCGGTGTGATGCTTAACCTTCAACATG 2760  
Db 2701 TTCTCACTTGTGTGAATCCCTTACCCTGGCATTCGGTGTGATGCTTAACCTTCAACATG 2760  
QY 2761 ATTCAAAATGATTTAAATGATCAGCAATTTTATGTACAGAAGAATATACATTATA 2820  
Db 2761 ATTCAAAATGATTTAAATGATCAGCAATTTTATGTACAGAAGAATATACATTATA 2820  
QY 2821 ATGCAATCCTCTGGGCTTTTACTCAAGGAAACGGCCATCCTCCCTAAATTTGACTTCG 2880  
Db 2821 ATGCAATCCTCTGGGCTTTTACTCAAGGAAACGGCCATCCTCCCTAAATTTGACTTCG 2880  
QY 2881 TTTTGTAGATGTCAGTGGCAGATGCCAAGAGGAGTATCATCAAAATGTCGATGCCGT 2940  
Db 2881 TTTTGTAGATGTCAGTGGCAGATGCCAAGAGGAGTATCATCAAAATGTCGATGCCGT 2940  
QY 2941 GTTTCGGAATGCTCTCACACCTACCAAAACAGGCGACCTTTCAGCAGAGAGATGATTTG 3000  
Db 2941 GTTTCGGAATGCTCTCACACCTACCAAAACAGGCGACCTTTCAGCAGAGAGATGATTTG 3000  
QY 3001 GGGCTACTCTCTCGCAGGCTCAGTGCAGAGTTCGTAGAGAACAAATTTAGTTTAAAG 3060  
Db 3001 GGGCTACTCTCTCGCAGGCTCAGTGCAGAGTTCGTAGAGAACAAATTTAGTTTAAAG 3060  
QY 3061 ACTTCATCCCTCCACCTATCCCTTAACAGCTCTAGATTACCAAAACAGATTAATTTAT 3120  
Db 3061 ACTTCATCCCTCCACCTATCCCTTAACAGCTCTAGATTACCAAAACAGATTAATTTAT 3120  
QY 3121 CACTAAAAGAAAATCTATTATCACTGCTCTACACAGCTTTTCTCTAGAACCGCTCT 3180  
Db 3121 CACTAAAAGAAAATCTATTATCACTGCTCTACACAGCTTTTCTCTAGAACCGCTCT 3180  
QY 3181 GCGTTTACTCTGTTTTCAAAAGGACTTTTGTAAAATCAAAATCATCTCTCACAAGGCAG 3240  
Db 3181 GCGTTTACTCTGTTTTCAAAAGGACTTTTGTAAAATCAAAATCATCTCTCACAAGGCAG 3240  
QY 3241 GAGGAGTGAATAAGAACTTTTATGGAGCTTATGATCGATCCAGGCTTCTCAGGCG 3300  
Db 3241 GAGGAGTGAATAAGAACTTTTATGGAGCTTATGATCGATCCAGGCTTCTCAGGCG 3300  
QY 3301 GCTTGTAGTGAATGCTGACCTGAAGTACAGTATATTCTTGTAAAATACATAAAACAAAGC 3360  
Db 3301 GCTTGTAGTGAATGCTGACCTGAAGTACAGTATATTCTTGTAAAATACATAAAACAAAGC 3360



QY 721 CATGAATATTGGAGGACATAAAGTGTGTGCCAGAAATGAACCTGGCAGGAATGC 780  
DB 721 CATGAATATTGGAGGACACATAAAGTGTGTGCCAGAAATGAACCTGGCAGGAATGC 780  
QY 781 ACCAGCTGTTCACATAGATCTAAATCAAACTCCTCAGACCACATTGCCACAATATTT 840  
DB 781 ACCAGCTGTTCACATAGATCTAAATCAAACTCCTCAGACCACATTGCCACAATATTT 840  
QY 841 CTTAAAGTAGGGAACCCCTATGGATAAGTGTCAAGCTGTTTATGTGAACCATGGATTC 900  
DB 841 CTTAAAGTAGGGAACCCCTATGGATAAGTGTCAAGCTGTTTATGTGAACCATGGATTC 900  
QY 901 GGGCTCACCTGGAAATAGAAAACAAGCACTCGAGGAGGCAACTACTTTTGAGATCAGT 960  
DB 901 GGGCTCACCTGGAAATAGAAAACAAGCACTCGAGGAGGCAACTACTTTTGAGATCAGT 960  
QY 961 ACCTATTCAACAACAAGAACTATGATACGGATTCTGTGTTTGTATCATCAGTGGCA 1020  
DB 961 ACCTATTCAACAACAAGAACTATGATACGGATTCTGTGTTTGTATCATCAGTGGCA 1020  
QY 1021 AGAAGCAGACCGGATACACTTGTCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080  
DB 1021 AGAAGCAGACCGGATACACTTGTCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080  
QY 1081 GTTACCATCGTAGGAAGGATTATAAATGCTACCAATTCAGTGAAGATTATGAATTT 1140  
DB 1081 GTTACCATCGTAGGAAGGATTATAAATGCTACCAATTCAGTGAAGATTATGAATTT 1140  
QY 1141 GACCAATATGAAGATTGTTGTTTCTGTCCAGCTTTAAAGCCTACCCACAATCAGATGT 1200  
DB 1141 GACCAATATGAAGATTGTTGTTTCTGTCCAGCTTTAAAGCCTACCCACAATCAGATGT 1200  
QY 1201 AGTGGACCTTCTCGAATAATCATTTCTGTGTGAGCAAAAGGCTGTGATACGGATAC 1260  
DB 1201 AGTGGACCTTCTCGAATAATCATTTCTGTGTGAGCAAAAGGCTGTGATACGGATAC 1260  
QY 1261 ASCATATCCAAGTTTGCATATAGCAACACCGAGGAGAAATATATATTCATTCAGAA 1320  
DB 1261 ASCATATCCAAGTTTGCATATAGCAACACCGAGGAGAAATATATATTCATTCAGAA 1320  
QY 1321 AATGATGATGCCCAATTTACCAAAATGTTTCAACGCTGAATAAAGAGAAACCTCAAGTG 1380  
DB 1321 AATGATGATGCCCAATTTACCAAAATGTTTCAACGCTGAATAAAGAGAAACCTCAAGTG 1380  
QY 1381 CTCGAGAAAGCATCGGCAAGTCAAGGCTGCTGTTTCTCGGATGGATACCCATCT 1440  
DB 1381 CTCGAGAAAGCATCGGCAAGTCAAGGCTGCTGTTTCTCGGATGGATACCCATCT 1440  
QY 1441 TGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCACTGCACAGAGATCACAGAGGA 1500  
DB 1441 TGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCACTGCACAGAGATCACAGAGGA 1500  
QY 1501 GTCGTGAATAGAAAGGCTAACGAAAGTGTGAGACAGTGGGTGCGAGCAGTACTCTA 1560  
DB 1501 GTCGTGAATAGAAAGGCTAACGAAAGTGTGAGACAGTGGGTGCGAGCAGTACTCTA 1560  
QY 1561 ACATGAGTGAAGCCATAAAGGTTCTCGTCAAGTGTGTCATACATAAATCCCTTGGC 1620  
DB 1561 ACATGAGTGAAGCCATAAAGGTTCTCGTCAAGTGTGTCATACATAAATCCCTTGGC 1620  
QY 1621 ACATCTGTGAGACGATCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAGACAAC 1680  
DB 1621 ACATCTGTGAGACGATCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAGACAAC 1680  
QY 1681 ATCTCATTTATGCAACAATGGTGTGTTGCTCTCTTCTCATTTGCTGTTTAAACCTGCTA 1740  
DB 1681 ATCTCATTTATGCAACAATGGTGTGTTGCTCTCTTCTCATTTGCTGTTTAAACCTGCTA 1740  
QY 1741 ATTTGTCACAAGTACAAAAGCAATTTAGTATGAAGCCAGCTACAGATGGTACAGGTG 1800  
DB 1741 ATTTGTCACAAGTACAAAAGCAATTTAGTATGAAGCCAGCTACAGATGGTACAGGTG 1800

QY 1801 ACCGGCTCCTCAGATAAATGAGTACTTCTACCTTGATTTCAGAGAATATGAATATGATCTC 1860  
DB 1801 ACCGGCTCCTCAGATAAATGAGTACTTCTACCTTGATTTCAGAGAATATGAATATGATCTC 1860  
QY 1861 AAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTCTTTT 1920  
DB 1861 AAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTCTTTT 1920  
QY 1921 GGAAGATGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980  
DB 1921 GGAAGATGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980  
QY 1981 GCCGTCAAAATGCTGAAAGAAAAGACACAGCTCTGAAAGAGAGGCACTCATGTCAAA 2040  
DB 1981 GCCGTCAAAATGCTGAAAGAAAAGACACAGCTCTGAAAGAGAGGCACTCATGTCAAA 2040  
QY 2041 CTCGAAGATGATGACCCAGCTGGGAAGCCACAGAAATATTTGTAACCTGCTGGGGCGTGC 2100  
DB 2041 CTCGAAGATGATGACCCAGCTGGGAAGCCACAGAAATATTTGTAACCTGCTGGGGCGTGC 2100  
QY 2101 ACATGTCAGGACCAATTTACTTGTATTTTGAATCTGTTGCTATGTTGATCTTCTCAAC 2160  
DB 2101 ACATGTCAGGACCAATTTACTTGTATTTTGAATCTGTTGCTATGTTGATCTTCTCAAC 2160  
QY 2161 TATCTAAGAAAGTAAAGAGAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
DB 2161 TATCTAAGAAAGTAAAGAGAAAATTTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
QY 2221 AATTTCAAGTTTACCACCTTCCCAATCAGATCCAAATTCAGCATGCCCTGGTTCAAGA 2280  
DB 2221 AATTTCAAGTTTACCACCTTCCCAATCAGATCCAAATTCAGCATGCCCTGGTTCAAGA 2280  
QY 2281 GAAGTTCCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGAATTCATTCAC 2340  
DB 2281 GAAGTTCCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGAATTCATTCAC 2340  
QY 2341 TCTGAAGATGAAATTTGAATGAAACCAAAAGGCTGGAAGAGAGAGGAGACTTGAAT 2400  
DB 2341 TCTGAAGATGAAATTTGAATGAAACCAAAAGGCTGGAAGAGAGAGGAGACTTGAAT 2400  
QY 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
DB 2401 GTGCTTACATTTGAAGATCTTCTTGTGCTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
QY 2461 CTGGAATTTAAGTGTGTGTTTACAGAGACCTGGCCGCGCAGGAACGTTGTCACCCAC 2520  
DB 2461 CTGGAATTTAAGTGTGTGTTTACAGAGACCTGGCCGCGCAGGAACGTTGTCACCCAC 2520  
QY 2521 GGAAGAGTGTGAAGATATGTGACTTTGGATTGCTCGAGATATCATGAGTATCCCAAC 2580  
DB 2521 GGAAGAGTGTGAAGATATGTGACTTTGGATTGCTCGAGATATCATGAGTATCCCAAC 2580  
QY 2581 TATGTTGTCAGGGCAATGCCGCTGCTGCTGTAATAATGGATGGCCCCCGAAGGCTGTTT 2640  
DB 2581 TATGTTGTCAGGGCAATGCCGCTGCTGCTGTAATAATGGATGGCCCCCGAAGGCTGTTT 2640  
QY 2641 GAAGGATCTACACCATTAAGAGTGTGTTGCTCATATGGAATATTACTGTGGGAATC 2700  
DB 2641 GAAGGATCTACACCATTAAGAGTGTGTTGCTCATATGGAATATTACTGTGGGAATC 2700  
QY 2701 TTTCTCACTTGTGTAATCCTTACCCTGGCATTCGGTGTGATGCTAATCTTCAAACTG 2760  
DB 2701 TTTCTCACTTGTGTAATCCTTACCCTGGCATTCGGTGTGATGCTAATCTTCAAACTG 2760  
QY 2761 ATTCAAAATGGAATTTAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATATA 2820  
DB 2761 ATTCAAAATGGAATTTAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATATA 2820  
QY 2821 ATGCAATCTGCTGGGCTTTGACTCAAGGAAGAGGCCATCTTCCCTTAATTTGACTTCG 2880  
DB 2821 ATGCAATCTGCTGGGCTTTGACTCAAGGAAGAGGCCATCTTCCCTTAATTTGACTTCG 2880  
QY 2881 TTTTGTAGGATGTCAGCTGGCAGATGCGAAGAACCGGATGTATCAGATGTGGATGGCCGT 2940

Db 2881 TTTTGGATGTCAGTGCAGATCAGAAAGCGATGATCAGAAATGTGATGCCCT 2940  
QY 2941 GTTTCGGATGTCCTCACACCTACCAACACACGGGACCTTTTCAGCAGAGAGATGATTTG 3000  
Db 2941 GTTTCGGATGTCCTCACACCTACCAACACACGGGACCTTTTCAGCAGAGAGATGATTTG 3000  
QY 3001 GGGCTACTCTCCCGAGGCTCAGTGCAGAGATTCGTAGAGAAACAAATTTAGTTTAAAG 3060  
Db 3001 GGGCTACTCTCCCGAGGCTCAGTGCAGAGATTCGTAGAGAAACAAATTTAGTTTAAAG 3060  
QY 3061 ACTTCATCCCTCCACCTATCCCTTAACAGGCTGTAGATTACCAAAACAAAGATTAATTTAT 3120  
Db 3061 ACTTCATCCCTCCACCTATCCCTTAACAGGCTGTAGATTACCAAAACAAAGATTAATTTAT 3120  
QY 3121 CACTAAAGAAATCTATTATCAACTGCTGTCCAGACACTTTTCTAGAGCCGTCT 3180  
Db 3121 CACTAAAGAAATCTATTATCAACTGCTGTCCAGACACTTTTCTAGAGCCGTCT 3180  
QY 3181 GGGTTTACTCTGTTTCAAGGACTTTTGTAAATCAAAATCAATCATCTGTCAAGGCAG 3240  
Db 3181 GGGTTTACTCTGTTTCAAGGACTTTTGTAAATCAAAATCAATCATCTGTCAAGGCAG 3240  
QY 3241 GAGGAGCTGATATGAACCTTTATTGGAGCATGATCTGCATCAAGGCCCTTCTCAGGCCG 3300  
Db 3241 GAGGAGCTGATATGAACCTTTATTGGAGCATGATCTGCATCAAGGCCCTTCTCAGGCCG 3300  
QY 3301 GCTTGAGTGAATGCTGACTGAAGTACAGTATATTTCTGTAAATCAATCAATCAATCAAGC 3360  
Db 3301 GCTTGAGTGAATGCTGACTGAAGTACAGTATATTTCTGTAAATCAATCAATCAATCAAGC 3360  
QY 3361 ATTTTCTAGGAGAGCTAATATGATTTTTTAAAGTCTATGTTTTTAAATCAATCAATCAAG 3420  
Db 3361 ATTTTCTAGGAGAGCTAATATGATTTTTTAAAGTCTATGTTTTTAAATCAATCAATCAAG 3420  
QY 3421 TTTTTCAGCTATTAGTATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480  
Db 3421 TTTTTCAGCTATTAGTATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA 3480  
QY 3481 AAAAAAAAAAAAAAAAAAAAA 3501  
Db 3481 AAAAAAAAAAAAAAAAAAAAA 3501

RESULT 5  
US-09-021-324-3  
Sequence 3, Application US/09021324  
Patent No. 5912133  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ImClone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/021.324  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/977.451  
FILING DATE: 1992-11-19  
PRIOR APPLICATION DATA:

Query Match 100.0%; Score 3501; DB 2; Length 3501;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGAGCGGCATCCGAGGGCTGGCGCGGCGCTGGGGACCCCGGGTCCCGAGGCCATG 60  
Db 1 CGAGCGGCATCCGAGGGCTGGCGCGGCGCTGGGGACCCCGGGTCCCGAGGCCATG 60  
QY 61 CCGCGGTTGGCGCGGACGCCGCTGCGTGTGTTTCTGCAATGATA 120  
Db 61 CCGCGGTTGGCGCGGACGCCGCTGCGTGTGTTTCTGCAATGATA 120  
QY 121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAATAG 180  
Db 121 TTTGGGACTATTACAAATCAAGATCTGCTGTGATCAAGTGTGTTTAAATCAATCAATAG 180  
QY 181 AACAAATGATTCATCAGTGGGAAGTCATCATCATATCCATGGTATCAGAAATCCCCGAA 240  
Db 181 AACAAATGATTCATCAGTGGGAAGTCATCATCATATCCATGGTATCAGAAATCCCCGAA 240  
QY 241 GACCTCGGGTGTGGTTGAGACCCAGAGCTCAGGACAGTGTACGAGCTGCCGTGTG 300  
Db 241 GACCTCGGGTGTGGTTGAGACCCAGAGCTCAGGACAGTGTACGAGCTGCCGTGTG 300



Db 241 GACCTCGGGTGTGCGTTGAGACCCCGAGAGCTCAGGGACAGTGTAACGAAGCTGCCGTGTG 300  
QY 301 GAAGTGGATGATATCTGCTCCATCACATGCTCAAGTGTCTGTCGATGCCCGCAGGGAACATTT 360  
Db 301 GAAGTGGATGATATCTGCTCCATCACATGCTCAAGTGTCTGTCGATGCCCGCAGGGAACATTT 360  
QY 361 TCCTGTCCTGGGTCTTTTAAAGCACAGCTCCCTGAAATGGCCAGCCACATTTTGTATTTACAA 420  
Db 361 TCCTGTCCTGGGTCTTTTAAAGCACAGCTCCCTGAAATGGCCAGCCACATTTTGTATTTACAA 420  
QY 421 AACAGAGGAGTGTCTTTCCATGCTCATTTTGAATAATGACAGAAACCCAGCTGAGATATAC 480  
Db 421 AACAGAGGAGTGTCTTTCCATGCTCATTTTGAATAATGACAGAAACCCAGCTGAGATATAC 480  
QY 481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTTACAGTGAGTATAAGA 540  
Db 481 CTACTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTTACAGTGAGTATAAGA 540  
QY 541 AATACCTCTCTTACACATTAAGAAGACCTTACTTTTAGAAAAATGGAACCCAGGAGGCC 600  
Db 541 AATACCTCTCTTACACATTAAGAAGACCTTACTTTTAGAAAAATGGAACCCAGGAGGCC 600  
QY 601 CTGGTCTGATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTGCGATTCA 660  
Db 601 CTGGTCTGATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTGCGATTCA 660  
QY 661 CAGGGGGAAGCTGTAAAGAAAGTCCAGCTGTCTTAAAGAGGAGGAAAGTGCTT 720  
Db 661 CAGGGGGAAGCTGTAAAGAAAGTCCAGCTGTCTTAAAGAGGAGGAAAGTGCTT 720  
QY 721 CATGAATTTTGGGACGGACATAAGTGTCTGCCAGAAATGAATGGGCGAGGAATGC 780  
Db 721 CATGAATTTTGGGACGGACATAAGTGTCTGCCAGAAATGAATGGGCGAGGAATGC 780  
QY 781 ACCAGCTCTTACAAATAGATCTAAATCAAACTCTCAGACCACATTTGCCACAATATTT 840  
Db 781 ACCAGCTCTTACAAATAGATCTAAATCAAACTCTCAGACCACATTTGCCACAATATTT 840  
QY 841 CTTAAAGTAGGGGAACCCCTATGATTAAGTGTCAAGCTGTTTCAATGTGAACCATGGATTC 900  
Db 841 CTTAAAGTAGGGGAACCCCTATGATTAAGTGTCAAGCTGTTTCAATGTGAACCATGGATTC 900  
QY 901 GGGCTCACCTGGGAATTAGAAAAACAAGCACTCGAGGAGGGCAACTACTTTTGAGATGAGT 960  
Db 901 GGGCTCACCTGGGAATTAGAAAAACAAGCACTCGAGGAGGGCAACTACTTTTGAGATGAGT 960  
QY 961 ACCTATTCAACAAACAGACTATGATACGGATCTGTTGCTTTGTATCATCAGTGGCA 1020  
Db 961 ACCTATTCAACAAACAGACTATGATACGGATCTGTTGCTTTGTATCATCAGTGGCA 1020  
QY 1021 AGAAACGACACCGGATACACTTGTCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080  
Db 1021 AGAAACGACACCGGATACACTTGTCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080  
QY 1081 GTTACCATCGTAGGAAGGATTTATAATGCTACCAATTCAAAGTGAAGATTTAGAAAT 1140  
Db 1081 GTTACCATCGTAGGAAGGATTTATAATGCTACCAATTCAAAGTGAAGATTTAGAAAT 1140  
QY 1141 GACCAATATGAAGAGTTTGTCTGTCAGGTTTAAAGCCCTACCCACAATCAGATGT 1200  
Db 1141 GACCAATATGAAGAGTTTGTCTGTCAGGTTTAAAGCCCTACCCACAATCAGATGT 1200  
QY 1201 ACGTGACCTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTCTTGTAAACGGATAC 1260  
Db 1201 ACGTGACCTCTCTCGAAATCATTTCTTGTGAGCAAAAGGCTCTTGTAAACGGATAC 1260  
QY 1261 AGCATATCCAAGTTTTCGAATCATAAGCACCGCCAGGGAATATATATTTCCATGAGAA 1320  
Db 1261 AGCATATCCAAGTTTTCGAATCATAAGCACCGCCAGGGAATATATATTTCCATGAGAA 1320  
QY 1321 AATGATGATGCCCAATTTACCAAAATGTTTCAACGCTGAATATAAGAAAGGAAACCTCAAGTG 1380  
Db 1321 AATGATGATGCCCAATTTACCAAAATGTTTCAACGCTGAATATAAGAAAGGAAACCTCAAGTG 1380

QY 1381 CTCGCAAGACATCGGCAAGTCAGCGCTCCTGTTTCTCGGATGGATACCCATTACCATCT 1440  
Db 1381 CTCGCAAGACATCGGCAAGTCAGCGCTCCTGTTTCTCGGATGGATACCCATTACCATCT 1440  
QY 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAAGA 1500  
Db 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAAGA 1500  
QY 1501 GTCTGGAATAGAAGGCTTAACAGAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
Db 1501 GTCTGGAATAGAAGGCTTAACAGAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
QY 1561 AACATGAGTGAAGCATATAAAGGGTCTCTGCTCAAGTCTGTGCATACAAATTCCTCTGGC 1620  
Db 1561 AACATGAGTGAAGCATATAAAGGGTCTCTGCTCAAGTCTGTGCATACAAATTCCTCTGGC 1620  
QY 1621 ACATCTTGTGAGAGGATCTCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGACAAC 1680  
Db 1621 ACATCTTGTGAGAGGATCTCTTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGACAAC 1680  
QY 1681 ATCTCATCTTATGCAACAATTTGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
Db 1681 ATCTCATCTTATGCAACAATTTGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
QY 1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Db 1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
QY 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTTGATTTTCAGAGAATATGAATATGATCTC 1860  
Db 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTTGATTTTCAGAGAATATGAATATGATCTC 1860  
QY 1861 AAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTCTTTT 1920  
Db 1861 AAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTCTTTT 1920  
QY 1921 GGAAGTGTATGACCAACAGCTTATGGATTTAGCAATAGCAAGAGTCTCAATCCAGTT 1980  
Db 1921 GGAAGTGTATGACCAACAGCTTATGGATTTAGCAATAGCAAGAGTCTCAATCCAGTT 1980  
QY 1981 GCCGTCAAAATGCTGAAAGAAAACAGACAGCTCTGAAAGAGAGGCACTCATGTCAAG 2040  
Db 1981 GCCGTCAAAATGCTGAAAGAAAACAGACAGCTCTGAAAGAGAGGCACTCATGTCAAG 2040  
QY 2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAAATATTTGTAACCTGCTGGGGCGTGC 2100  
Db 2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAAATATTTGTAACCTGCTGGGGCGTGC 2100  
QY 2101 ACATGTCAGGACCAATTTACTTGTATTTTGAATACTGTTGATGTTCTTCTCAAC 2160  
Db 2101 ACATGTCAGGACCAATTTACTTGTATTTTGAATACTGTTGATGTTCTTCTCAAC 2160  
QY 2161 TATCTAAAGAGTAAAGAGAAAATTTTACAGGACTTTGGACAGAGATTTTCAAGGAACAC 2220  
Db 2161 TATCTAAAGAGTAAAGAGAAAATTTTACAGGACTTTGGACAGAGATTTTCAAGGAACAC 2220  
QY 2221 AATTTCAAGTTTACCCTCTTTCCATTCATCCAAATTTCCAGCATGCCCTGGTTCAAGA 2280  
Db 2221 AATTTCAAGTTTACCCTCTTTCCATTCATCCAAATTTCCAGCATGCCCTGGTTCAAGA 2280  
QY 2281 GAAGTTTCAGATACACCCGAGCTCGGATCAATCTCAGGGCTTTCATGGGAATTCATTAC 2340  
Db 2281 GAAGTTTCAGATACACCCGAGCTCGGATCAATCTCAGGGCTTTCATGGGAATTCATTAC 2340  
QY 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAAGGCTGGAAGAGAGGAGGACTTTGAAT 2400  
Db 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAAAGGCTGGAAGAGAGGAGGACTTTGAAT 2400  
QY 2401 GTGCTTACATTTGAAGATCTTCTTTTGTCTTTCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
Db 2401 GTGCTTACATTTGAAGATCTTCTTTTGTCTTTCATATCAAGTTGCCAAAGGAATGGAATTT 2460

Qy	2461	CTGGAATTTAAAGTCGTGTGTTTCACAGAGACCTGGCCGCCAGGAACCTGCTTGTTCACCCAC	2520
Db	2461	CTGGAATTTAAAGTCGTGTGTTTCACAGAGACCTGGCCGCCAGGAACCTGCTTGTTCACCCAC	2520
Qy	2521	GGGAAGTGGTGAAGATATGTGTGACTTTGGATTGGCTCGAGATATCATGATGATTCACCAAC	2580
Db	2521	GGGAAGTGGTGAAGATATGTGTGACTTTGGATTGGCTCGAGATATCATGATGATTCACCAAC	2580
Qy	2581	TATGTTGTCAAGGGCAATGCCCGCTCTGCCCTGTAATAATGGATGGCCCCCGAAGACCCCTGTTT	2640
Db	2581	TATGTTGTCAAGGGCAATGCCCGCTCTGCCCTGTAATAATGGATGGCCCCCGAAGACCCCTGTTT	2640
Qy	2641	GAAGGCATCTACACCATTAAGAGTGTGTCTGCATATGGAATATTACTGTGGGAAATC	2700
Db	2641	GAAGGCATCTACACCATTAAGAGTGTGTCTGCATATGGAATATTACTGTGGGAAATC	2700
Qy	2701	TTCTCACTGGTGTGAATCCTTACCCTGGCATTCGGGTGATCCTCACTTCTACAACCTG	2760
Db	2701	TTCTCACTGGTGTGAATCCTTACCCTGGCATTCGGGTGATCCTCACTTCTACAACCTG	2760
Qy	2761	ATTCAAAATGGATTTAAATGGATCAGCCATTTATGCTACAGAAGAAATATACATTATA	2820
Db	2761	ATTCAAAATGGATTTAAATGGATCAGCCATTTATGCTACAGAAGAAATATACATTATA	2820
Qy	2821	ATGCAATCCCTGCTGGGCTTTTGACTCAAGAAACGGCCATCCTTCCCTAAATTTGACTTCG	2880
Db	2821	ATGCAATCCCTGCTGGGCTTTTGACTCAAGAAACGGCCATCCTTCCCTAAATTTGACTTCG	2880
Qy	2881	TTTTTAGATGTGAGCTGGCAGATGCAAGAGACCGATGTATCAGATGTGGATGGCCGT	2940
Db	2881	TTTTTAGATGTGAGCTGGCAGATGCAAGAGACCGATGTATCAGATGTGGATGGCCGT	2940
Qy	2941	GTTTCGGAATGCTCTCAACCTACCAAAACAGGGACCTTTTCAGCAGAGATGGGATTTG	3000
Db	2941	GTTTCGGAATGCTCTCAACCTACCAAAACAGGGACCTTTTCAGCAGAGATGGGATTTG	3000
Qy	3001	GGGCTACTCTCCGCGAGGCTCAGTGCAGAGATTCGTAGAGAACAAATTTAGTTTAAAG	3060
Db	3001	GGGCTACTCTCCGCGAGGCTCAGTGCAGAGATTCGTAGAGAACAAATTTAGTTTAAAG	3060
Qy	3061	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGGATTAATTTTCAT	3120
Db	3061	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAGGATTAATTTTCAT	3120
Qy	3121	CACATAAAGAAATCTATTATCACTGCTGCTTCCACAGACTTTTCTCTAGAAGCCGCTCT	3180
Db	3121	CACATAAAGAAATCTATTATCACTGCTGCTTCCACAGACTTTTCTCTAGAAGCCGCTCT	3180
Qy	3181	CGGTTTACTCTGTGTTTCAAAGGACCTTTGTAAATCAAAATCATCTGTCCACAGGCAG	3240
Db	3181	CGGTTTACTCTGTGTTTCAAAGGACCTTTGTAAATCAAAATCATCTGTCCACAGGCAG	3240
Qy	3241	GAGGAGCTGATAATGAACCTTTATTGGAGCATGTATCTGATCCAGGCCCTCTCAGGCGG	3300
Db	3241	GAGGAGCTGATAATGAACCTTTATTGGAGCATGTATCTGATCCAGGCCCTCTCAGGCGG	3300
Qy	3301	GCTTGCAGTGAATGTGTACCTGGAAGTACAGTATATCTTGTGAAATACATAAACAAGC	3360
Db	3301	GCTTGCAGTGAATGTGTACCTGGAAGTACAGTATATCTTGTGAAATACATAAACAAGC	3360
Qy	3361	ATTTTGTCAAGGAGCAATATGATTTTTTAAAGTCTATCTTTTTTAAATTAATATGTA	3420
Db	3361	ATTTTGTCAAGGAGCAATATGATTTTTTAAAGTCTATCTTTTTTAAATTAATATGTA	3420
Qy	3421	TTTTTTCAGCTATTTAGTGATATATTTTATGGGTGGGAATAAAATTTCTTACTACAGAAA	3480
Db	3421	TTTTTTCAGCTATTTAGTGATATATTTTATGGGTGGGAATAAAATTTCTTACTACAGAAA	3480
Qy	3481	AAAAAAAAAAAAAAAAAAAAA 3501	
Db	3481	AAAAAAAAAAAAAAAAAAAAA 3501	

QY 361 TCCGTGCTCTGGTCTTTAAGCACAGCTCCCTGAATTCGAGCCACATTTTGAATTTACAA 420  
DB 361 TCCGTGCTCTGGTCTTTAAGCACAGCTCCCTGAATTCGAGCCACATTTTGAATTTACAA 420  
QY 421 AACAGGAGTGTGTTTCCATGTCATTTGAAATGACAGAAACCCAAAGCTGGAGATAC 480  
DB 421 AACAGGAGTGTGTTTCCATGTCATTTGAAATGACAGAAACCCAAAGCTGGAGATAC 480  
QY 481 CTACTTTTATTCAGAGTGAAGCTACCAATACACAATATGTTTACAGTGAATTAAGA 540  
DB 481 CTACTTTTATTCAGAGTGAAGCTACCAATACACAATATGTTTACAGTGAATTAAGA 540  
QY 541 AATACCCCTGCTTTACACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACCAAGCAGCC 600  
DB 541 AATACCCCTGCTTTACACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACCAAGCAGCC 600  
QY 601 CTGGTCTGATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTCGGATTC 660  
DB 601 CTGGTCTGATATCTCAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTCGGATTC 660  
QY 661 CAGGGGAAAGCTGTAAGAAGAAAGCTCCAGCTGTTGTTAAAAAGAGGAAAGTGCCTT 720  
DB 661 CAGGGGAAAGCTGTAAGAAGAAAGCTCCAGCTGTTGTTAAAAAGAGGAAAGTGCCTT 720  
QY 721 CATGAATTTATTTGGACGGACATAAGGTGCTGTCAGAAATGAACCTGGGCAGGGAATGC 780  
DB 721 CATGAATTTATTTGGACGGACATAAGGTGCTGTCAGAAATGAACCTGGGCAGGGAATGC 780  
QY 781 ACCAGGCTGTTCCAAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCACAATATTT 840  
DB 781 ACCAGGCTGTTCCAAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCACAATATTT 840  
QY 841 CTTAAAGTAGGGAACCCCTATGGAATAGGTGCAAGCTGTTTCATGTGAACCATGATTC 900  
DB 841 CTTAAAGTAGGGAACCCCTATGGAATAGGTGCAAGCTGTTTCATGTGAACCATGATTC 900  
QY 901 GGGCTCACCTGGGAATTAGAAAAACAAGCACTCGAGGAGGGCAACTACTTTTGAGATGAGT 960  
DB 901 GGGCTCACCTGGGAATTAGAAAAACAAGCACTCGAGGAGGGCAACTACTTTTGAGATGAGT 960  
QY 961 ACCTATTCAACAAACAGAACTATGATPACGGATCTGTTGCTTTTGTATCATCAGTGGCA 1020  
DB 961 ACCTATTCAACAAACAGAACTATGATPACGGATCTGTTGCTTTTGTATCATCAGTGGCA 1020  
QY 1021 AGAAAGCACCGGATACACTTCTCTTCAAGCATCCCAAGTCAATCAGCTTTC 1080  
DB 1021 AGAAAGCACCGGATACACTTCTCTTCAAGCATCCCAAGTCAATCAGCTTTC 1080  
QY 1081 GTTACCATCTAGGAAAGGATTTATAATGCTTACCAATTCAGTGAAGATTTATGAATTT 1140  
DB 1081 GTTACCATCTAGGAAAGGATTTATAATGCTTACCAATTCAGTGAAGATTTATGAATTT 1140  
QY 1141 GACCAATATGAAGATTTTGTGTTTCTGTCAGGTTTAAAGCCCTPACCCACAAATCAGATGT 1200  
DB 1141 GACCAATATGAAGATTTTGTGTTTCTGTCAGGTTTAAAGCCCTPACCCACAAATCAGATGT 1200  
QY 1201 AGTGGACCTCTCTCGAAATCATTTCTTGTGACCAAAAGGCTTGTATACGGATAC 1260  
DB 1201 AGTGGACCTCTCTCGAAATCATTTCTTGTGACCAAAAGGCTTGTATACGGATAC 1260  
QY 1261 AGCATATCCAAGTTTGTCAATCATAGCACACCGAGGAGATATATATTTCCATGCAGAA 1320  
DB 1261 AGCATATCCAAGTTTGTCAATCATAGCACACCGAGGAGATATATATTTCCATGCAGAA 1320  
QY 1321 AATGATGATGCCAATTTTACCAAAATGTTTACGCTGAATATAGAAGGAAACCTCAAGTG 1380  
DB 1321 AATGATGATGCCAATTTTACCAAAATGTTTACGCTGAATATAGAAGGAAACCTCAAGTG 1380  
QY 1381 CTCGAGAGCATCGGCAAGTCAAGGCTGCTTCTCGGATGGATACCCATTTACCATCT 1440  
DB 1381 CTCGAGAGCATCGGCAAGTCAAGGCTGCTTCTCGGATGGATACCCATTTACCATCT 1440

QY 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAAGGA 1500  
DB 1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATCACAGAAGGA 1500  
QY 1501 GTCTGGAATAGAAGGCTAACAGAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
DB 1501 GTCTGGAATAGAAGGCTAACAGAAAAGTGTTCGACAGTGGGTGTCGAGCAGTACTCTA 1560  
QY 1561 AACATGAGTCAAGCCATAAAGGGTTCCTGGTCAAGTGTCTGTGCATACAATTCCTTGGC 1620  
DB 1561 AACATGAGTCAAGCCATAAAGGGTTCCTGGTCAAGTGTCTGTGCATACAATTCCTTGGC 1620  
QY 1621 ACATCTTGTGAGAGTCTCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGACAAC 1680  
DB 1621 ACATCTTGTGAGAGTCTCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGACAAC 1680  
QY 1681 ATCTCATCTATGCAACAATGTTGTTGCTCTCTCTTCATPTGTCGTTTTAACCCCTGCTA 1740  
DB 1681 ATCTCATCTATGCAACAATGTTGTTGCTCTCTCTCTTCATPTGTCGTTTTAACCCCTGCTA 1740  
QY 1741 ATTTGTCAAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
DB 1741 ATTTGTCAAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
QY 1801 ACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAAATATGATCTC 1860  
DB 1801 ACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAAATATGATCTC 1860  
QY 1861 AAATGGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAAGGTACTTAGGATCAGGTGCTTTT 1920  
DB 1861 AAATGGGAGTTTCCAGAGAAAATTTAGAGTTTGGGAAGGTACTTAGGATCAGGTGCTTTT 1920  
QY 1921 GGAAGTGTATGACGCAACAGCTTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
DB 1921 GGAAGTGTATGACGCAACAGCTTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
QY 1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAAG 2040  
DB 1981 GCCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAAG 2040  
QY 2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAAATATGTTGAACCTGCTGGGGCGGTGC 2100  
DB 2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAAATATGTTGAACCTGCTGGGGCGGTGC 2100  
QY 2101 ACAGTGTGAGGACCAATTTTACTGATTTTGAATCTGTTGCTATGTTGATCTTCTCAAC 2160  
DB 2101 ACAGTGTGAGGACCAATTTTACTGATTTTGAATCTGTTGCTATGTTGATCTTCTCAAC 2160  
QY 2161 TATCTAAGAAAGTAAAAGAGAAAATTTTCACAGGACTTTGGACAGAGATTTTCAAGGAACAC 2220  
DB 2161 TATCTAAGAAAGTAAAAGAGAAAATTTTCACAGGACTTTGGACAGAGATTTTCAAGGAACAC 2220  
QY 2221 AATTTCAAGTTTTTACCCACTTTTCCAATCATCTCAATTTCCAGCATGCTGTTCAAGA 2280  
DB 2221 AATTTCAAGTTTTTACCCACTTTTCCAATCATCTCAATTTCCAGCATGCTGTTCAAGA 2280  
QY 2281 GAAGTTCAGATACACCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCAAC 2340  
DB 2281 GAAGTTCAGATACACCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCAAC 2340  
QY 2341 TCTGAAGATCAAAATGAATATGAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
DB 2341 TCTGAAGATCAAAATGAATATGAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
QY 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTCATATCAAGTTTGCCTCAAGGAATTTT 2460  
DB 2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTCATATCAAGTTTGCCTCAAGGAATTTT 2460  
QY 2461 CTGGAATTTAAGTGTGTTCACAGACCTGGCCGCCAGGAACTGCTTGTGACCCAC 2520  
DB 2461 CTGGAATTTAAGTGTGTTCACAGACCTGGCCGCCAGGAACTGCTTGTGACCCAC 2520  
QY 2521 GGGAAAGTGTGAGATATGTGACTTTGGATTGGCTCGAGATATCATGATGATTTCCAAC 2580

2521	Db		GGGAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTCAATCCAAC	2580
2581	Qy		TATGTTGTTCAGGGCAATGCCCGTCTGCTGTAAAAATGGATGGCCCGCGAAAGCCGTGTTT	2640
2581	Db		TATGTTGTTCAGGGCAATGCCCGTCTGCTGTAAAAATGGATGGCCCGCGAAAGCCGTGTT	2640
2641	Qy		GAAGGCATCTACACCATTAAAGAGTGAATGCTCGGTCAATATGGAATATTAAGTGTGGGAATC	2700
2641	Db		GAAGGCATCTACACCATTAAAGAGTGAATGCTCGGTCAATATGGAATATTAAGTGTGGGAATC	2700
2701	Qy		TTCTCACATGGTGTGAATCCTTTACCTCGCATTTCCGGTTGATGCTAACTTCTACAAACATG	2760
2701	Db		TTCTCACATGGTGTGAATCCTTTACCTCGCATTTCCGGTTGATGCTAACTTCTACAAACATG	2760
2761	Qy		ATTCAAAATGGATTTAAAAATGGATCAGCCATTATTAAGTCTACAGAAGAAATATACATTATA	2820
2761	Db		ATTCAAAATGGATTTAAAAATGGATCAGCCATTATTAAGTCTACAGAAGAAATATACATTATA	2820
2821	Qy		ATGCAATCCTGCTGGGCTTTTGACTCAAGGAACGGCCATCTCTCCCTCAATTTGACTTCG	2880
2821	Db		ATGCAATCCTGCTGGGCTTTTGACTCAAGGAACGGCCATCTCTCCCTCAATTTGACTTCG	2880
2881	Qy		TTTTTAGTAGTGTACGTGGCAGATGCAGAAGCGATGTATCAGAATGTGGATGCCCGT	2940
2881	Db		TTTTTAGTAGTGTACGTGGCAGATGCAGAAGCGATGTATCAGAATGTGGATGCCCGT	2940
2941	Qy		GTTTCGGAATGCTCTACACCTTACCAAAACAGCGGACCTTTTCAGCAGAGAGATGGATTTC	3000
2941	Db		GTTTCGGAATGCTCTACACCTTACCAAAACAGCGGACCTTTTCAGCAGAGAGATGGATTTC	3000
3001	Qy		GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGAACCAATTTAGTTTTAAGG	3060
3001	Db		GGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGAACCAATTTAGTTTTAAGG	3060
3061	Qy		ACTTCAATCCCTCACCTATGCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTTTCAT	3120
3061	Db		ACTTCAATCCCTCACCTATGCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTTTCAT	3120
3121	Qy		CACATAAAGAAAATCTATTATCAACTGCTCTTCACAGACATTTTCTCTAGAACGCCGTCT	3180
3121	Db		CACATAAAGAAAATCTATTATCAACTGCTCTTCACAGACATTTTCTCTAGAACGCCGTCT	3180
3181	Qy		GCCTTTTACTCTGTTTTCAAAAGGACTTTTGTAAAAATCAATCATCTGTCAACAAGGCAG	3240
3181	Db		GCCTTTTACTCTGTTTTCAAAAGGACTTTTGTAAAAATCAATCATCTGTCAACAAGGCAG	3240
3241	Qy		GAGGAGCTGATATGAACCTTATTGGAGCATTCGATCTGCATCCAAGGCCCTTCTCAGGCCG	3300
3241	Db		GAGGAGCTGATATGAACCTTATTGGAGCATTCGATCTGCATCCAAGGCCCTTCTCAGGCCG	3300
3301	Qy		GCCTTAGTGAATTTGCTACCTGAAGTACAGTATATCTTCTTAATACATAAAAAACAAGC	3360
3301	Db		GCCTTAGTGAATTTGCTACCTGAAGTACAGTATATCTTCTTAATACATAAAAAACAAGC	3360
3361	Qy		ATTTTGTCTAGGAGAAGCTAATATGATTTTTTTTAACTATGTTTTTAAATAAATATGTAAA	3420
3361	Db		ATTTTGTCTAGGAGAAGCTAATATGATTTTTTTTAACTATGTTTTTAAATAAATATGTAAA	3420
3421	Qy		TTTTTTCAGCTATTATTAGTGATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA	3480
3421	Db		TTTTTTCAGCTATTATTAGTGATATATTTTATGGTGGGAATAAAATTTCTACTACAGAAAA	3480
3481	Qy		AAAAAAAAAAAAAAAAAAAAA 3501	
3481	Db		AAAAAAAAAAAAAAAAAAAAA 3501	

RESULT 7  
PCT-US92-09893-3  
; Sequence 3, Application PC/TUS9209893  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.

TITLE OF INVENTION: TORIOPOTENT HEMATOPOIETIC STEM CELL  
 TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
 NUMBER OF SEQUENCES: 10  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Imclone Systems Incorporated  
 STREET: 180 Varick Street  
 CITY: New York  
 STATE: New York  
 COUNTRY: U.S.A.  
 ZIP: 10014  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US92/09893  
 FILING DATE: 19921116  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Feit, Irving N.  
 REGISTRATION NUMBER: 28,601  
 REFERENCE/DOCKET NUMBER: LEM-3-7PT  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-645-1405  
 TELEFAX: 212-645-2054  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3501 base pairs  
 TYPE: NUCLEIC ACID  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: CDNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: N-terminal  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 58..3039  
 FEATURE:  
 NAME/KEY: mat\_peptide  
 LOCATION: 139..3036  
 FEATURE:  
 NAME/KEY: sig\_peptide  
 LOCATION: 58..138  
 PCT-US92-09893-3

Query Match	100.0%	Score	3501;	DB	5;	Length	3501;
Best Local Similarity	100.0%;	Pred.	No. 0;				
Matches	3501;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	CGAGGGCGCATCCGAGGGCTGGGGCCGGCCCTGGGGACCCCGGGCTCCGGAGGCCCATG	60				
Db	1	CGAGGGCGCATCCGAGGGCTGGGGCCGGCCCTGGGGACCCCGGGCTCCGGAGGCCCATG	60				
Qy	61	CCGGCGTTGGCGCGGACGGCGGGCACCGTGCCTGCTCGTGTGTTTTCTGCAATGATA	120				
Db	61	CCGGCGTTGGCGCGGACGGCGGGCACCGTGCCTGCTCGTGTGTTTTCTGCAATGATA	120				
Qy	121	TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTCGTGTTTAAATCAATCATAAAG	180				
Db	121	TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTCGTGTTTAAATCAATCATAAAG	180				
Qy	181	AACAATGATTTCATCAGTGGGGAAGTCATCATCATATCCATGGTATCAGAATCCCGCGAA	240				
Db	181	AACAATGATTTCATCAGTGGGGAAGTCATCATCATATCCATGGTATCAGAATCCCGCGAA	240				
Qy	241	GACCTCGGGTGTGGCTTGAGACCCGAGAGCTCAGGGACAGTGTACGAAGCTCCCGCTGTG	300				
Db	241	GACCTCGGGTGTGGCTTGAGACCCGAGAGCTCAGGGACAGTGTACGAAGCTCCCGCTGTG	300				
Qy	301	GAAGTGGATGTATCTGCTTCCATCACACTGCAAGTGTGGTCGATGCCCCAGGGCAACATT	360				

Db 301 GAAGTGGATGTATCTGCTTCCATCAGCTGCAAGTCTGCTCGATGCCCGCAGGAACATT 360  
Qy 361 TCCTGTCTCTGGTCTTTAAGCACAGCTCCCTGAATTGCCAGCCACATTTTGTATTTACAA 420  
Db 361 TCCTGTCTCTGGTCTTTAAGCACAGCTCCCTGAANTGCCAGCCACATTTTGTATTTACAA 420  
Qy 421 AACAGAGGAGTGTGTTCATGTGTCATTTTGAATAATGACAGAAACCCCAAGCTGGAGAATAC 480  
Db 421 AACAGAGGAGTGTGTTCATGTGTCATTTTGAATAATGACAGAAACCCCAAGCTGGAGAATAC 480  
Qy 481 CTACTTTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTGTACAGTGAATATAAGA 540  
Db 481 CTACTTTTTTATTCAGAGTGAAGCTACCAATTTACACAATATTTGTACAGTGAATATAAGA 540  
Qy 541 AATACCCCTGCTTTACACATTTAAGAGAGCTTACTTTTAGAAAAATGGAACACAGGACGCC 600  
Db 541 AATACCCCTGCTTTACACATTTAAGAGAGCTTACTTTTAGAAAAATGGAACACAGGACGCC 600  
Qy 601 CTGGTCTCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTTCGATTCA 660  
Db 601 CTGGTCTCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTTCGATTCA 660  
Qy 661 CAGGGGGAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGAAAGTGCTT 720  
Db 661 CAGGGGGAAGCTGTAAGAAGAAAGTCCAGCTGTTGTTTAAAGAGGAGAAAGTGCTT 720  
Qy 721 CATGAATTTATTTGGACCGGACATAAGGTGCTGTGCCAGAAATGAAGTGGCAGGGAATGC 780  
Db 721 CATGAATTTATTTGGACCGGACATAAGGTGCTGTGCCAGAAATGAAGTGGCAGGGAATGC 780  
Qy 781 ACCAGGCTGTTTCCAAATAGATCTAAATCAAACTCTCAGACACACATTTGCCACAAATTTAT 840  
Db 781 ACCAGGCTGTTTCCAAATAGATCTAAATCAAACTCTCAGACACACATTTGCCACAAATTTAT 840  
Qy 841 CTTAAAGTAGGGGAACCCCTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCATGGATT 900  
Db 841 CTTAAAGTAGGGGAACCCCTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCATGGATT 900  
Qy 901 GGGCTCACCTGGGAATTAGAAAACAAAGACATCGAGGAGGCACTACTTTGAGATGAT 960  
Db 901 GGGCTCACCTGGGAATTAGAAAACAAAGACATCGAGGAGGCACTACTTTGAGATGAT 960  
Qy 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTGTCTGTTTGTATCATCAGTGGCA 1020  
Db 961 ACCTATTCAACAAACAGAACTATGATACGGATCTGTTTGTCTGTTTGTATCATCAGTGGCA 1020  
Qy 1021 AGAAGACACCGGATACACTGTGTCCTTCCAAAGCATCCAGTCCAGTCAATCAGCTTTG 1080  
Db 1021 AGAAGACACCGGATACACTGTGTCCTTCCAAAGCATCCAGTCCAGTCAATCAGCTTTG 1080  
Qy 1081 GTTACCATCGTAGGAAGGGATTTAATGCTACCAATTTCAAGTGAAGATTATGAAT 1140  
Db 1081 GTTACCATCGTAGGAAGGGATTTAATGCTACCAATTTCAAGTGAAGATTATGAAT 1140  
Qy 1141 GACCAATATGAAGAGTGTGTTTCTGTGTCAGTGTAAAGCCTACCCCAAAATCAGATGT 1200  
Db 1141 GACCAATATGAAGAGTGTGTTTCTGTGTCAGTGTAAAGCCTACCCCAAAATCAGATGT 1200  
Qy 1201 ACGTGGACCTCTCTCGAAAATCATTTCTGTGAGCAAAAAGGGCTTTGATAACGGATAC 1260  
Db 1201 ACGTGGACCTCTCTCGAAAATCATTTCTGTGAGCAAAAAGGGCTTTGATAACGGATAC 1260  
Qy 1261 AGCATATCCAAAGTTTGAATCATAGCACAGCCAGGAGGAATATATTTCCATCGAGAA 1320  
Db 1261 AGCATATCCAAAGTTTGAATCATAGCACAGCCAGGAGGAATATATTTCCATCGAGAA 1320  
Qy 1321 AATGATGATGCCCAATTTACCAAAATGTTCAAGCTGTAATAAGAAGGAACCTCAAGTG 1380  
Db 1321 AATGATGATGCCCAATTTACCAAAATGTTCAAGCTGTAATAAGAAGGAACCTCAAGTG 1380  
Qy 1381 CTCGCAAGACATCGGCAAGTCAGGCGTCTGTTTCTCGGATGGATACCCATTACCATCT 1440  
Db 1381 CTCGCAAGACATCGGCAAGTCAGGCGTCTGTTTCTCGGATGGATACCCATTACCATCT 1440

Qy 1441 TGACCTTGGAGAGTGTCTTCAGACAACTCTCCCACTGCACAGAGAGATCACAGAAGGA 1500  
Db 1441 TGACCTTGGAGAGTGTCTTCAGACAACTCTCCCACTGCACAGAGAGATCACAGAAGGA 1500  
Qy 1501 GTCTGGAATAGAAGGCTAACAGAAAAGTGTGAGACAGTGGGTGTGAGCAGTACTCTA 1560  
Db 1501 GTCTGGAATAGAAGGCTAACAGAAAAGTGTGAGACAGTGGGTGTGAGCAGTACTCTA 1560  
Qy 1561 ACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTCTGATCAATCAATTCCTTGGC 1620  
Db 1561 ACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTCTGATCAATCAATTCCTTGGC 1620  
Qy 1621 ACATCTTGTGAGAGTCTCTTTTAAACTCTCCAGGCCCTTCCCTTTTCCATCCAGAAC 1680  
Db 1621 ACATCTTGTGAGAGTCTCTTTTAAACTCTCCAGGCCCTTCCCTTTTCCATCCAGAAC 1680  
Qy 1681 ATCTCATTTTATCAACAATTTGGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
Db 1681 ATCTCATTTTATCAACAATTTGGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
Qy 1741 ATTTGTCAAGTACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Db 1741 ATTTGTCAAGTACAAAAGCAATTTAGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Qy 1801 ACCGGCTCTCAGATATGAGTACTTCTACGTTGATTTTTCAGAGAAATATGATATGATCTC 1860  
Db 1801 ACCGGCTCTCAGATATGAGTACTTCTACGTTGATTTTTCAGAGAAATATGATATGATCTC 1860  
Qy 1861 AAATGGAGTTCCTCAAGAGAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT 1920  
Db 1861 AAATGGAGTTCCTCAAGAGAAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT 1920  
Qy 1921 GGAAGTGTATGAAACACAGCTTATGGAATTTAGCAAAAACAGAGTCTCAATCCAGGTT 1980  
Db 1921 GGAAGTGTATGAAACACAGCTTATGGAATTTAGCAAAAACAGAGTCTCAATCCAGGTT 1980  
Qy 1981 GCGCTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAAG 2040  
Db 1981 GCGCTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAAG 2040  
Qy 2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGAATTTCTGAACCTGCTGGGGCGTGC 2100  
Db 2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGAATTTCTGAACCTGCTGGGGCGTGC 2100  
Qy 2101 ACATCTCAGGACCAATTTACTTGAATTTTGAATCTGTTGCTATGGTATCTTCTCAAC 2160  
Db 2101 ACATCTCAGGACCAATTTACTTGAATTTTGAATCTGTTGCTATGGTATCTTCTCAAC 2160  
Qy 2161 TATCTAAGAGTAAAGAGAAAATTTTACAGGACTTGGACAGAGATTTTCAGGAACAC 2220  
Db 2161 TATCTAAGAGTAAAGAGAAAATTTTACAGGACTTGGACAGAGATTTTCAGGAACAC 2220  
Qy 2221 AATTTTCAGTTTACCCCACTTTCCCAATCACATCCCAATTTCCAGCATGCTGTTCAAGA 2280  
Db 2221 AATTTTCAGTTTACCCCACTTTCCCAATCACATCCCAATTTCCAGCATGCTGTTCAAGA 2280  
Qy 2281 GAAGTTCAGATACACCCGACCTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340  
Db 2281 GAAGTTCAGATACACCCGACCTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340  
Qy 2341 TCTGAAGATGAATTTGAATATGAAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
Db 2341 TCTGAAGATGAATTTGAATATGAAAACCAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
Qy 2401 GTGCTTACATTTGAAGATCTTCTTTCGATATCAAGTTCGCCAAAGGAATGGAAATTT 2460  
Db 2401 GTGCTTACATTTGAAGATCTTCTTTCGATATCAAGTTCGCCAAAGGAATGGAAATTT 2460  
Qy 2461 CTGGAATTTAAGTCTGTGTTTTCAGAGACCTTGGCGCCAGGAACTGCTTGTACCCAC 2520  
Db 2461 CTGGAATTTAAGTCTGTGTTTTCAGAGACCTTGGCGCCAGGAACTGCTTGTACCCAC 2520



QY 421 AACAGAGAGTGTGTTCCATGCTCATTTGAAATACAGAAACCCCAAGCTGGAGATAC 480  
DB 421 AACAGAGAGTGTGTTCCATGCTCATTTGAAATACAGAAACCCCAAGCTGGAGATAC 480  
QY 481 CTACTTTTATTCAGAGTGAAGCTACCAATATACAAATATGTTTACAGTGAATAGA 540  
DB 481 CTACTTTTATTCAGAGTGAAGCTACCAATATACAAATATGTTTACAGTGAATAGA 540  
QY 541 AATACCTGCTTTACACATTAAGNAGACCTTACTTTAGAAAAATGGAACCCAGGAGCC 600  
DB 541 AATACCTGCTTTACACATTAAGNAGACCTTACTTTAGAAAAATGGAACCCAGGAGCC 600  
QY 601 CTGGTCTGCATATCTGAGAGCGTTCAGAGCGGATCGTGAATGGTGTTCGATTCA 660  
DB 601 CTGGTCTGCATATCTGAGAGCGTTCAGAGCGGATCGTGAATGGTGTTCGATTCA 660  
QY 661 CAGGGGAAAGCTGTAAGAAAGAAAGTCCAGCTGTTGTTTAAAAAGGAGAAAAAGTCTT 720  
DB 661 CAGGGGAAAGCTGTAAGAAAGAAAGTCCAGCTGTTGTTTAAAAAGGAGAAAAAGTCTT 720  
QY 721 CATGAATATTTGGGACGGACATAAGGTGCTGTCGCGCAGAAATGAACCTGGGAGGAATGC 780  
DB 721 CATGAATATTTGGGACGGACATAAGGTGCTGTCGCGCAGAAATGAACCTGGGAGGAATGC 780  
QY 781 ACCAGGCTGTTCACATPAGATCTAAATCAAACTCCTCAGACCACATTCGCCACATTTATTT 840  
DB 781 ACCAGGCTGTTCACATPAGATCTAAATCAAACTCCTCAGACCACATTCGCCACATTTATTT 840  
QY 841 CTTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCAAAAGCTGTTCAATGTAACCATTTATTT 900  
DB 841 CTTTAAAGTAGGGAAACCCCTTATGGATAAGGTGCAAAAGCTGTTCAATGTAACCATTTATTT 900  
QY 901 GGCTCACCTGGAAATAGAAAACAAAGACTCGAGAGGGCAACTACTTTGAGATGAGT 960  
DB 901 GGCTCACCTGGAAATAGAAAACAAAGACTCGAGAGGGCAACTACTTTGAGATGAGT 960  
QY 961 ACCTATTCAACAACAGAACTATGATACGGATCTGTTGCTTTGTTGATCATCAGTGC 1020  
DB 961 ACCTATTCAACAACAGAACTATGATACGGATCTGTTGCTTTGTTGATCATCAGTGC 1020  
QY 1021 AGAAACGACACCGGATCTACTACTTGTCTCTTCAAAAGCATCCAGTCAATCAGCTTTG 1080  
DB 1021 AGAAACGACACCGGATCTACTACTTGTCTCTTCAAAAGCATCCAGTCAATCAGCTTTG 1080  
QY 1081 GTTACCATCTAGAAAGGATTTATTAATGCTACCAATTTCAAGTGAAGATTAAGAAAT 1140  
DB 1081 GTTACCATCTAGAAAGGATTTATTAATGCTACCAATTTCAAGTGAAGATTAAGAAAT 1140  
QY 1141 GACCAATATGAAGAGTGTCTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200  
DB 1141 GACCAATATGAAGAGTGTCTTTCTGTCAGGTTTAAAGCCTACCCACAAATCAGATGT 1200  
QY 1201 ACSTGGACCTTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTGTATACGGATAC 1260  
DB 1201 ACSTGGACCTTCTCGAAATCATTTCTTGTGAGCAAAAGGCTTGTATACGGATAC 1260  
QY 1261 AGCATATCCAAGTTTTCATCATATAGCACCAGCAGGAGAAATATATATTCATTCAGAA 1320  
DB 1261 AGCATATCCAAGTTTTCATCATATAGCACCAGCAGGAGAAATATATATTCATTCAGAA 1320  
QY 1321 AATGATGATGCCCAATTTACCAAAATGTTTCAACGCTGAATATAGAAGGAACCTCAAGTG 1380  
DB 1321 AATGATGATGCCCAATTTACCAAAATGTTTCAACGCTGAATATAGAAGGAACCTCAAGTG 1380  
QY 1381 CTGCGAGAAGCATCGCAAGTCAAGGCTCTGTTTCTCGGATGGATACCCATTAACCATCT 1440  
DB 1381 CTGCGAGAAGCATCGCAAGTCAAGGCTCTGTTTCTCGGATGGATACCCATTAACCATCT 1440  
QY 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCACTGCACAGAAAGATCACAGAAGGA 1500  
DB 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCACTGCACAGAAAGATCACAGAAGGA 1500

QY 1501 GTCTGGATAGAAAGGCTAACAGAAAAAGTGTGACAGTGGTGTGACAGTACTCTA 1560  
DB 1501 GTCTGGATAGAAAGGCTAACAGAAAAAGTGTGACAGTGGTGTGACAGTACTCTA 1560  
QY 1561 AACATGAGTGAAGCCATAAAAGGTTCTGTCTCAAGTGTGTGCATCAATTCCTTGGC 1620  
DB 1561 AACATGAGTGAAGCCATAAAAGGTTCTGTCTCAAGTGTGTGCATCAATTCCTTGGC 1620  
QY 1621 ACATCTTGTGAGAGATCTTTTAACTCTCCAGGCCCTTCCCTTTTCAATCAAGAAC 1680  
DB 1621 ACATCTTGTGAGAGATCTTTTAACTCTCCAGGCCCTTCCCTTTTCAATCAAGAAC 1680  
QY 1681 ATCTCATTTATGCAACAATTTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
DB 1681 ATCTCATTTATGCAACAATTTGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
QY 1741 ATTTGTCAACAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGATACAGT 1800  
DB 1741 ATTTGTCAACAGTACAAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGATACAGT 1800  
QY 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTACAGAGATATGATATGATCTC 1860  
DB 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTACAGAGATATGATATGATCTC 1860  
QY 1861 AAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGTCTCTTT 1920  
DB 1861 AAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGTCTCTTT 1920  
QY 1921 GGAAGTGTATGAAAGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGT 1980  
DB 1921 GGAAGTGTATGAAAGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGT 1980  
QY 1981 GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCA 2040  
DB 1981 GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCA 2040  
QY 2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAAATTTGTAACCTCTCTGGGCGTGC 2100  
DB 2041 CTCAGATGATGACCCAGCTGGGAAGCCACAGAAATTTGTAACCTCTCTGGGCGTGC 2100  
QY 2101 ACATGTCAAGCAATTTTACTTGTATTTTGAATCTGTTGATGTTGATCTCTCAAC 2160  
DB 2101 ACATGTCAAGCAATTTTACTTGTATTTTGAATCTGTTGATGTTGATCTCTCAAC 2160  
QY 2161 TATCTAAGAGTAAAGAGAAAAATTTTACAGGACTTTGGACAGAGATTTTCAAGGAAC 2220  
DB 2161 TATCTAAGAGTAAAGAGAAAAATTTTACAGGACTTTGGACAGAGATTTTCAAGGAAC 2220  
QY 2221 AATTTAGTTTTACCCTTCCCAATCAGTCAATTTCCAGCATCCCAATTTCCAGTCC 2280  
DB 2221 AATTTAGTTTTACCCTTCCCAATCAGTCAATTTCCAGCATCCCAATTTCCAGTCC 2280  
QY 2281 GAATTCAGATACACCGGACTCGGATCAATCTCAGGCTTTCATGGAAATTCATTTCC 2340  
DB 2281 GAATTCAGATACACCGGACTCGGATCAATCTCAGGCTTTCATGGAAATTCATTTCC 2340  
QY 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAGAGAGGACTTTGA 2400  
DB 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAGAGAGGACTTTGA 2400  
QY 2401 GTGCTTACATTTGAAGATCTTCTTGTCTTTCATATCAAGTTGCAAGGATGGAATTT 2460  
DB 2401 GTGCTTACATTTGAAGATCTTCTTGTCTTTCATATCAAGTTGCAAGGATGGAATTT 2460  
QY 2461 CTGGAATTTAAGTCTGTTTACAGAGACCTGGCCGCGAGGAGTCTGTTGTCACCCAC 2520  
DB 2461 CTGGAATTTAAGTCTGTTTACAGAGACCTGGCCGCGAGGAGTCTGTTGTCACCCAC 2520  
QY 2521 GGGAAAGTGTGAAAGATATGAGTCTTGGATTTGGTCTCGAGATATCATGAGTATTC 2580  
DB 2521 GGGAAAGTGTGAAAGATATGAGTCTTGGATTTGGTCTCGAGATATCATGAGTATTC 2580  
QY 2581 TATGTTGTCAAGGCAATGCCCTGCTGCTGTAATGAGTGGCCCCCGAAGGCTGTTT 2640



[illegible]



2701	Db	TTCTCACTTGGTGTGAATCCCTTACCCTGGCATTTCCGGTTGATGCTTAACCTCTTACAACCTG	2716
2761	Qy	ATTCAAATCGATTAAAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA	2820
2761	Db	ATTCAAATGGATTAAAAATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATA	2820
2821	Qy	ATGCAATCCTGCTGGGCTTTTGACTCAAGGAAACAGGCCATCCTTCCCTAATTGCACTTCG	2880
2821	Db	ATGCAATCCTGCTGGGCTTTTGACTCAAGGAAACAGGCCATCCTTCCCTAATTGCACTTCG	2880
2881	Qy	TTTTTTAGGATGTCAGCTGGCAGATGCAGAGAAGCGATGTATCAGATGTGGATGGCCGT	2940
2881	Db	TTTTTTAGGATGTCAGCTGGCAGATGCAGAGAAGCGATGTATCAGATGTGGATGGCCGT	2940
2941	Qy	GTWTCGGAATGTCCTCACACCTACCACAAACAGGCGACCTTTTCAGCAGAGAGATGGAATTG	3000
2941	Db	GTWTCGGAATGTCCTCACACCTACCACAAACAGGCGACCTTTTCAGCAGAGAGATGGAATTG	3000
3001	Qy	GGGCTACTCTCTCCGCAAGGCTCAGGTCGAAGATTCGTPAGAGAACAAATTTAGTTTTAAGG	3060
3001	Db	GGGCTACTCTCTCCGCAAGGCTCAGGTCGAAGATTCGTPAGAGAACAAATTTAGTTTTAAGG	3060
3061	Qy	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTATTTTCAT	3120
3061	Db	ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTATTTTCAT	3120
3121	Qy	CACATAAAGAAATCTATTATCAACTGCTGCTTCACACAGACTTTTCTCAGAACCCGTCT	3180
3121	Db	CACATAAAGAAATCTATTATCAACTGCTGCTTCACACAGACTTTTCTCAGAACCCGTCT	3180
3181	Qy	CGGTTTACTCTGTGTTTTCAAAGGACCTTTTGTAATAATCAATCATCCTGTCCACAGGCAG	3240
3181	Db	CGGTTTACTCTGTGTTTTCAAAGGACCTTTTGTAATAATCAATCATCCTGTCCACAGGCAG	3240
3241	Qy	GAGGAGCTGATAATGAACCTTTATTGGAGCATGTGATCTGCAATCCAGGCCCTTCTCAGGCCG	3300
3241	Db	GAGGAGCTGATAATGAACCTTTATTGGAGCATGTGATCTGCAATCCAGGCCCTTCTCAGGCCG	3300
3301	Qy	GCTTGAGTGAAATGTGTACCTGGAAGTACAGTATATCTTGTAATAATACATAAACAAGC	3360
3301	Db	GCTTGAGTGAAATGTGTACCTGGAAGTACAGTATATCTTGTAATAATACATAAACAAGC	3360
3361	Qy	ATTTTGTCTAAGGAGAAGCTAATATGATTTTTTAAAGTCTATGTTTTTAAATAATATGTAAA	3420
3361	Db	ATTTTGTCTAAGGAGAAGCTAATATGATTTTTTAAAGTCTATGTTTTTAAATAATATGTAAA	3420
3421	Qy	TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAAATTTCTACACAG	3475
3421	Db	TTTTTCAGCTATTTAGTGATATATTTTATGGTGGGAATAAAATTTCTACACAG	3475

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

|||||  
721 CATGAATATTTGGGACGGACATAAGGTGCTGTGCCAGAAATGAACCTGGCGAGGAATGC 780  
QY  
781 ACCAGGCTGTTCAACATAGATCTAAATCAAACTCCCTCAGACCACATGTCACACATATATTT 840  
Db  
781 ACCAGGCTGTTCAACATAGATCTAAATCAAACTCCCTCAGACCACATGTCACACATATATTT 840  
QY  
841 CTTAAAGTAGGGGAACCCCTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCATGGATTC 900  
Db  
841 CTTAAAGTAGGGGAACCCCTTATGGATAAGGTGCAAGCTGTTTCATGTGAACCATGGATTC 900  
QY  
901 GGGCTCACCTGGGAATTAGAAAAACAAGCACTCGAGGAGGGCAACTACTTTTGAGATGAGT 960  
Db  
901 GGGCTCACCTGGGAATTAGAAAAACAAGCACTCGAGGAGGGCAACTACTTTTGAGATGAGT 960  
QY  
961 ACCTATTCAACAACAGAACTATGATACGGATTCTGTGTTGCTTTTGTATCATCAGTGGCA 1020  
Db  
961 ACCTATTCAACAACAGAACTATGATACGGATTCTGTGTTGCTTTTGTATCATCAGTGGCA 1020  
QY  
1021 AGAAACGACACCGGATACACTTGTCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080  
Db  
1021 AGAAACGACACCGGATACACTTGTCTCTTCAAGCATCCAGTCAATCAGCTTTG 1080  
QY  
1081 GTTACCATCGTAGGAAAGGATTTATAATGCTACCAATTCAAAGTGAAGATTTATGAAAT 1140  
Db  
1081 GTTACCATCGTAGGAAAGGATTTATAATGCTACCAATTCAAAGTGAAGATTTATGAAAT 1140  
QY  
1141 GACCAATATGAAGAGTTTGTGTTCTGTCAGGTTTAAAGCCATCCCAACAAATCAGATGT 1200  
Db  
1141 GACCAATATGAAGAGTTTGTGTTCTGTCAGGTTTAAAGCCATCCCAACAAATCAGATGT 1200  
QY  
1201 ACGTGACCTCTCTCGAAATCATTTCTGTTGAGCAAAAGGCTCTTGATAACGGATAC 1260  
Db  
1201 ACGTGACCTCTCTCGAAATCATTTCTGTTGAGCAAAAGGCTCTTGATAACGGATAC 1260  
QY  
1261 AGCATATCCAAAGTTTGCATATCAATAGCACACAGCCAGGAGAAATATATATCCATGCAGAA 1320  
Db  
1261 AGCATATCCAAAGTTTGCATATCAATAGCACACAGCCAGGAGAAATATATATCCATGCAGAA 1320  
QY  
1321 AATGATGATGCCAATTTACCAAAATGTTCAACGCTGAATATAAGAGGAAACCTCAAGTG 1380  
Db  
1321 AATGATGATGCCAATTTACCAAAATGTTCAACGCTGAATATAAGAGGAAACCTCAAGTG 1380  
QY  
1381 CTCGAGAGAGCATCGCAAGTCAGGCGTCTGTTCTCGAGTGAATACCCATACCATCT 1440  
Db  
1381 CTCGAGAGAGCATCGCAAGTCAGGCGTCTGTTCTCGAGTGAATACCCATACCATCT 1440  
QY  
1441 TGGACCTGGAAGAAGTGTTCAGACAAAGTCTCCCACTGCACAGAGAGATCACAGAAGGA 1500  
Db  
1441 TGGACCTGGAAGAAGTGTTCAGACAAAGTCTCCCACTGCACAGAGAGATCACAGAAGGA 1500  
QY  
1501 GTCTGGAATAGAAAGCTAACGAAAGTGTGAGACAGTGGGTGTCGAGCAGTACTCTA 1560  
Db  
1501 GTCTGGAATAGAAAGCTAACGAAAGTGTGAGACAGTGGGTGTCGAGCAGTACTCTA 1560  
QY  
1561 AACATGAGTGAAGCCATAAAGGTTCTCGTGGTCAAGTGTGTCATACAAATCCCTTGGC 1620  
Db  
1561 AACATGAGTGAAGCCATAAAGGTTCTCGTGGTCAAGTGTGTCATACAAATCCCTTGGC 1620  
QY  
1621 ACATCTGTGAGACGATCCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAAAGACAAC 1680  
Db  
1621 ACATCTGTGAGACGATCCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAAAGACAAC 1680  
QY  
1681 ATCTCATTTATGCAACAATGGTGTGTTCTCCTCTTCATGTCGTTTAAACCTGCTA 1740  
Db  
1681 ATCTCATTTATGCAACAATGGTGTGTTCTCCTCTTCATGTCGTTTAAACCTGCTA 1740  
QY  
1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Db  
1741 ATTTGTCACAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
QY  
1801 ACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAAATATCAATATGATCTC 1860  
|||||

Db  
1801 ACCGGCTCCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGATCTC 1860  
QY  
1861 AAATGGGAGTTTCCAAAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
Db  
1861 AAATGGGAGTTTCCAAAGAAAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
QY  
1921 GGAAGCTGATGAACGCAACAGCTTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980  
Db  
1921 GGAAGCTGATGAACGCAACAGCTTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGTT 1980  
QY  
1981 GCGCTCAAAATGCTGAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTFCAGAA 2040  
Db  
1981 GCGCTCAAAATGCTGAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTFCAGAA 2040  
QY  
2041 CTCAGATGATGATGCCAGCTGGGAAGCCACAGAAATATTGTGAACCTGCTGGGGCGTGC 2100  
Db  
2041 CTCAGATGATGATGCCAGCTGGGAAGCCACAGAAATATTGTGAACCTGCTGGGGCGTGC 2100  
QY  
2101 ACAGTGTGAGACCAATTTACTTGAATTTTCAATCTGTTGATGTTGATGTTCTCAAC 2160  
Db  
2101 ACAGTGTGAGACCAATTTACTTGAATTTTCAATCTGTTGATGTTGATGTTCTCAAC 2160  
QY  
2161 TATCTAAGAAAGTAAAGAGAAAAATTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
Db  
2161 TATCTAAGAAAGTAAAGAGAAAAATTTACAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
QY  
2221 AATTTCAAGTTTACCCCACTTTCCATCATCATCCAAATCCAGCATGCCCTGGTTCAAGA 2280  
Db  
2221 AATTTCAAGTTTACCCCACTTTCCATCATCATCCAAATCCAGCATGCCCTGGTTCAAGA 2280  
QY  
2281 GAAGTTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCCAC 2340  
Db  
2281 GAAGTTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCCAC 2340  
QY  
2341 TCTGAAGATGAATTTGAATATGAAAGCAAAAGGCTGGAAGAGAGAGGAGACTTTGAAT 2400  
Db  
2341 TCTGAAGATGAATTTGAATATGAAAGCAAAAGGCTGGAAGAGAGAGGAGACTTTGAAT 2400  
QY  
2401 GTGCTTACATTTGAAGATCTTCTTTGCTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
Db  
2401 GTGCTTACATTTGAAGATCTTCTTTGCTTGCATATCAAGTTGCCAAAGGAATGGAATTT 2460  
QY  
2461 CTGGAATTTAAGTCGTGTTTCAGAGACCTGCGCCAGAGAACTGCTTGTCCACCCAC 2520  
Db  
2461 CTGGAATTTAAGTCGTGTTTCAGAGACCTGCGCCAGAGAACTGCTTGTCCACCCAC 2520  
QY  
2521 GGAAGAGTGTGAAGATATGTGACTTTGGAATGCTCGAGATATCATGAGTCAATTCACAC 2580  
Db  
2521 GGAAGAGTGTGAAGATATGTGACTTTGGAATGCTCGAGATATCATGAGTCAATTCACAC 2580  
QY  
2581 TATGTTGTCAGGGCAATGCCGCTGCTGCTGTAATAATGGATGGCCCGGAAAGCTGTTT 2640  
Db  
2581 TATGTTGTCAGGGCAATGCCGCTGCTGCTGTAATAATGGATGGCCCGGAAAGCTGTTT 2640  
QY  
2641 GAAGGCACTACACCAATTAAGAGTGTGCTGCTGATGGAATATTACTGTGGGAATC 2700  
Db  
2641 GAAGGCACTACACCAATTAAGAGTGTGCTGCTGATGGAATATTACTGTGGGAATC 2700  
QY  
2701 TTCTCACTTGTGATCTTACCCCTGGCATTCGGGTGATGCTTAATCTCTACAAACTG 2760  
Db  
2701 TTCTCACTTGTGATCTTACCCCTGGCATTCGGGTGATGCTTAATCTCTACAAACTG 2760  
QY  
2761 ATTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATTATA 2820  
Db  
2761 ATTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATTATA 2820  
QY  
2821 ATGCAATCTGCTGGGCTTTTGAATCAAGGAAAGGCGCATCTTCCCTTAATTTGACTTCG 2880  
Db  
2821 ATGCAATCTGCTGGGCTTTTGAATCAAGGAAAGGCGCATCTTCCCTTAATTTGACTTCG 2880  
QY  
2881 TTTTATGAGGATGTCAGTGGCAGATGCAAGAGAGCGGATGATCAGAAATGTCGATGGCGGT 2940  
Db  
2881 TTTTATGAGGATGTCAGTGGCAGATGCAAGAGAGCGGATGATCAGAAATGTCGATGGCGGT 2940



QY 901 GGCTCACCTGGGAATTTAGAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT 960  
DB 901 GGCTCACCTGGGAATTTAGAAAACAAAGCACTCGAGGAGGGCAACTACTTTGAGATGAGT 960  
QY 961 ACCTATTCAACAACAGAACTATGATACGGATTCGTTTGTGATCATCATCAGTGGCA 1020  
DB 961 ACCTATTCAACAACAGAACTATGATACGGATTCGTTTGTGATCATCATCAGTGGCA 1020  
QY 1021 AGAAACGACACCGGATACTACACTTGTCTCTTCAAGGATCCAGCTCAATCAGCTTGG 1080  
DB 1021 AGAAACGACACCGGATACTACACTTGTCTCTTCAAGGATCCAGCTCAATCAGCTTGG 1080  
QY 1081 GTTACCATCGTAGGAAAGGATTTATAAATGCTACCAATTTCAAGTGAAGATTTATGAAT 1140  
DB 1081 GTTACCATCGTAGGAAAGGATTTATAAATGCTACCAATTTCAAGTGAAGATTTATGAAT 1140  
QY 1141 GACCAATATGAAGATTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200  
DB 1141 GACCAATATGAAGATTTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200  
QY 1201 ACCTGACCTTCTCTCGAAATCATTTCCCTGTGTCAGCAAAAGGCTTGTGTAACGGATAC 1260  
DB 1201 ACCTGACCTTCTCTCGAAATCATTTCCCTGTGTCAGCAAAAGGCTTGTGTAACGGATAC 1260  
QY 1261 AGCATATCCAAGTTTTCGAATCATTAAGCACAGCAGGAGAAATATATATTCATTCAGAA 1320  
DB 1261 AGCATATCCAAGTTTTCGAATCATTAAGCACAGCAGGAGAAATATATATTCATTCAGAA 1320  
QY 1321 AATGATGATGCCAAATTTACCAAAATGTTTCCAGCTGTAATATAAGAAAGAAACCTTCAAGT 1380  
DB 1321 AATGATGATGCCAAATTTACCAAAATGTTTCCAGCTGTAATATAAGAAAGAAACCTTCAAGT 1380  
QY 1381 CTCGCAAGACATCGCAAGTCTAGGCGTCTCTGTTTCTCGATGATGATACCATACATCT 1440  
DB 1381 CTCGCAAGACATCGCAAGTCTAGGCGTCTCTGTTTCTCGATGATGATACCATACATCT 1440  
QY 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCACTGTCAGCAAGAGATCACAGAAGGA 1500  
DB 1441 TGGACCTGGAAGAGTGTTCAGACAAAGTCTCCCACTGTCAGCAAGAGATCACAGAAGGA 1500  
QY 1501 GTCTGGAATAGAAAGGCTAACGAAAGTGTGTCAGAGTGGGTGTCGAGCAGTACTCTA 1560  
DB 1501 GTCTGGAATAGAAAGGCTAACGAAAGTGTGTCAGAGTGGGTGTCGAGCAGTACTCTA 1560  
QY 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTCTGATCAATATCCCTTGGC 1620  
DB 1561 AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTCTGATCAATATCCCTTGGC 1620  
QY 1621 ACATCTGTGAGACGATCTCTTTTAACTCTCCAGGCGCCCTTCCCTTTCATCCAGACAAC 1680  
DB 1621 ACATCTGTGAGACGATCTCTTTTAACTCTCCAGGCGCCCTTCCCTTTCATCCAGACAAC 1680  
QY 1681 ATCTCATTTATGCAACAATTTGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
DB 1681 ATCTCATTTATGCAACAATTTGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1740  
QY 1741 ATTGTGTCAGAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGATGATGATG 1800  
DB 1741 ATTGTGTCAGAGTACAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGATGATGATG 1800  
QY 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTTCAGAGAAATATGATATCTC 1860  
DB 1801 ACCGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTTCAGAGAAATATGATATCTC 1860  
QY 1861 AAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGTTTT 1920  
DB 1861 AAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGTTTT 1920  
QY 1921 GGAAGAGTATGAACGCAACAGCTTATGGAATTTAGCAAAACAGAGGTCTCAATCCAGGTT 1980  
DB 1921 GGAAGAGTATGAACGCAACAGCTTATGGAATTTAGCAAAACAGAGGTCTCAATCCAGGTT 1980

QY 1981 GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
DB 1981 GCCGTCAAAATGCTGAAAGAAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAA 2040  
QY 2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGAAATATGTGAACCTGCTGGGGGCGTGC 2100  
DB 2041 CTCAGATGATGACCCAGCTGGGAAGCCAGAGAAATATGTGAACCTGCTGGGGGCGTGC 2100  
QY 2101 ACCTGTGACAGCACTTACTTGAATTTTGAATACTGTTGCTATGTTGATCTTCTCAAC 2160  
DB 2101 ACCTGTGACAGCACTTACTTGAATTTTGAATACTGTTGCTATGTTGATCTTCTCAAC 2160  
QY 2161 TATCTAAGAACTAAGAGAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
DB 2161 TATCTAAGAACTAAGAGAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACAC 2220  
QY 2221 AATTTAGTTTTTACCCCACTTTCCAAATCACAATCCAAATTCAGCATGCCGTGGTTCAAG 2280  
DB 2221 AATTTAGTTTTTACCCCACTTTCCAAATCACAATTCAGCATGCCGTGGTTCAAG 2280  
QY 2281 GAAGTTCAGATACACCCGGACTCGGATCAAAATCTCAGGCTTTCAGGAAATTCATTTCAAC 2340  
DB 2281 GAAGTTCAGATACACCCGGACTCGGATCAAAATCTCAGGCTTTCAGGAAATTCATTTCAAC 2340  
QY 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAGAGAGGAGACTTGAAT 2400  
DB 2341 TCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAGAGAGGAGACTTGAAT 2400  
QY 2401 GTGCTTACATTTGAAGATCTTCTTGTCTTGCATATCAAGTTGCCAAAGGAAATGGAAT 2460  
DB 2401 GTGCTTACATTTGAAGATCTTCTTGTCTTGCATATCAAGTTGCCAAAGGAAATGGAAT 2460  
QY 2461 CTGGAATTTAAGTCTGTGTTTACAGAGACCTGGCCGCGAGGAACTGCTTGTCAACCCAC 2520  
DB 2461 CTGGAATTTAAGTCTGTGTTTACAGAGACCTGGCCGCGAGGAACTGCTTGTCAACCCAC 2520  
QY 2521 GGGAAAGTGGTGAAGATATGAGTCTTGGATTTGGCTCGAGATATCATGAGTATTCACAC 2580  
DB 2521 GGGAAAGTGGTGAAGATATGAGTCTTGGATTTGGCTCGAGATATCATGAGTATTCACAC 2580  
QY 2581 TATGTTCTGAGGGCAATGCCGCTGCTGCTGTAATGATGGCTGGCCGCGAGGAACTGTTT 2640  
DB 2581 TATGTTCTGAGGGCAATGCCGCTGCTGCTGTAATGATGGCTGGCCGCGAGGAACTGTTT 2640  
QY 2641 GAAGGATCTACACCATTTAGAGTATGCTGTCATATGGAATATTTACTGTGGGAATC 2700  
DB 2641 GAAGGATCTACACCATTTAGAGTATGCTGTCATATGGAATATTTACTGTGGGAATC 2700  
QY 2701 TTCTCACTTGGTGAATCTTACCTGGCAATTCGGTGTGATGCTTAACCTTCTACAACTG 2760  
DB 2701 TTCTCACTTGGTGAATCTTACCTGGCAATTCGGTGTGATGCTTAACCTTCTACAACTG 2760  
QY 2761 ATTCAAAATGGATTTAAATGGATCAGCAATTTTATGCTACAGAGAAATATACATATA 2820  
DB 2761 ATTCAAAATGGATTTAAATGGATCAGCAATTTTATGCTACAGAGAAATATACATATA 2820  
QY 2821 ATGCAATCTGCTGGGCTTTTACTCAAGGAAACGGCCATCTTCCCTTAATTTGACTTCG 2880  
DB 2821 ATGCAATCTGCTGGGCTTTTACTCAAGGAAACGGCCATCTTCCCTTAATTTGACTTCG 2880  
QY 2881 TTTTGTAGGATGTCAGCTGGGAGATGCAGAGAAAGGATGATCAGATGTCGATGCGGCT 2940  
DB 2881 TTTTGTAGGATGTCAGCTGGGAGATGCAGAGAAAGGATGATCAGATGTCGATGCGGCT 2940  
QY 2941 GTTTCGGAATGTCCTCACACTACCAAAACAGGCGACTTTTACAGAGAGATGATGATG 3000  
DB 2941 GTTTCGGAATGTCCTCACACTACCAAAACAGGCGACTTTTACAGAGAGATGATGATG 3000  
QY 3001 GGGCTACTCTCTCGCAGGCTCAGGTGGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG 3060  
DB 3001 GGGCTACTCTCTCGCAGGCTCAGGTGGAAGATTCGTAGAGGAACAATTTAGTTTTAAGG 3060  
QY 3061 ACTTCATCCCTCCACTTATCCCTAACAGGCTGTAGATTTACCAAAACAGATTAATTCAT 3120

Matches 3475; Conservative		0;	Mismatches	1;	Indels	Q;	Gaps	0;
Qy	1	CGAGGGGATCGAGGGCTGGCCGGCGCCCTGGGGACCCGGGCTCCGAGGCCATG	60					
Db	1	CGAGGGGATCGAGGGCTGGCCGGCGCCCTGGGGACCCGGGCTCCGAGGCCATG	60					
Qy	61	CCGGCGTGGCGCGACGCGGCGACCGTGGCGCTGCTCGTTGTTTTTCTGCAATGATA	120					
Db	61	CCGGCGTGGCGCGACGCGGCGACCGTGGCGCTGCTCGTTGTTTTTCTGCAATGATA	120					
Qy	121	TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTTTAATCAATATAAG	180					
Db	121	TTTGGGACTATTACAAATCAAGATCTGCCTGTGATCAAGTGTGTTTTTAATCAATATAAG	180					
Qy	181	AACAATGATTCATGAGTGGGAAGTGCATCATATCCCATGGTATCAGAAATCCCGGAA	240					
Db	181	AACAATGATTCATGAGTGGGAAGTGCATCATATCCCATGGTATCAGAAATCCCGGAA	240					
Qy	241	GACCTGGGTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCGCGTGTG	300					
Db	241	GACCTGGGTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCGCGTGTG	300					
Qy	301	GAAGTGATGATCTGCTTCCATCACACTGCAAGTCTGGTCGATCCCCAGGGAACATT	360					
Db	301	GAAGTGATGATCTGCTTCCATCACACTGCAAGTCTGGTCGATCCCCAGGGAACATT	360					
Qy	361	TCCTGCTCTGGGTCTTTAAGCAGAGCTCCCTGAATGCCAGCCACATTTTGATTTACAA	420					
Db	361	TCCTGCTCTGGGTCTTTAAGCAGAGCTCCCTGAATGCCAGCCACATTTTGATTTACAA	420					
Qy	421	AACAGAGGAGTTGTTCCATGTCATTTTCAAAATGACAGAAACCAAGCTGAGAAATC	480					
Db	421	AACAGAGGAGTTGTTCCATGTCATTTTCAAAATGACAGAAACCAAGCTGAGAAATC	480					
Qy	481	CTACTTTTATTCAGAGTGAAGCTACCAATTACACAATATGTTTACAGTGCATATAAGA	540					
Db	481	CTACTTTTATTCAGAGTGAAGCTACCAATTACACAATATGTTTACAGTGCATATAAGA	540					
Qy	541	AATACCTGCTTTACACATTAAGAAGACCTTACTTTAGAAAATGGAACACAGGACGCC	600					
Db	541	AATACCTGCTTTACACATTAAGAAGACCTTACTTTAGAAAATGGAACACAGGACGCC	600					
Qy	601	CTGCTCTGCATATCTGAGACGGTTCAGAGCCGATCGTGGAAATGGTGCTTTCCGATTCA	660					
Db	601	CTGCTCTGCATATCTGAGACGGTTCAGAGCCGATCGTGGAAATGGTGCTTTCCGATTCA	660					
Qy	661	CAGGGGAAAGCTCTAAAGAAAGTCCAGCTGTTGTTTAAAAGGAGGAAAAGTGCTT	720					
Db	661	CAGGGGAAAGCTCTAAAGAAAGTCCAGCTGTTGTTTAAAAGGAGGAAAAGTGCTT	720					
Qy	721	CATCAATATTTGGGACGGACATAAAGTGTGTGCCAGAAATGAACCTGGCAGGGAATGC	780					
Db	721	CATCAATATTTGGGACGGACATAAAGTGTGTGCCAGAAATGAACCTGGCAGGGAATGC	780					
Qy	781	ACCAGGCTGTTACAAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCACAATATTTT	840					
Db	781	ACCAGGCTGTTACAAATAGATCTAAATCAAACTCCTCAGACCACATTTGCCACAATATTTT	840					
Qy	841	CTTAAAGTAGGGGAACCTTATGATTAAGGTGCAAACTCTTCATGTGAACCATGATTTC	900					
Db	841	CTTAAAGTAGGGGAACCTTATGATTAAGGTGCAAACTCTTCATGTGAACCATGATTTC	900					
Qy	901	GGGCTCACCTGGGAATTTAGAAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATGAGT	960					
Db	901	GGGCTCACCTGGGAATTTAGAAAACAAAGCACTCGAGGAGGCAACTACTTTGAGATGAGT	960					
Qy	961	ACCTATTCAACAAACAGAAGCTATGATACGATTCCTGTTTCTTATCATCAGTGGCA	1020					
Db	961	ACCTATTCAACAAACAGAAGCTATGATACGATTCCTGTTTCTTATCATCAGTGGCA	1020					
Qy	1021	AGAAACGACACCGGATACTACACTTGTTCCTCTTCAAGCATCCAGTCATCAGCTTTC	1080					
Db	1021	AGAAACGACACCGGATACTACACTTGTTCCTCTTCAAGCATCCAGTCATCAGCTTTC	1080					

```

Query Match          99.2%; Score 3474.4; DB 5; Length 3476;
Best Local Similarity 100.0%; Pred. No. 0;

```



1081 GTTACCATCGTAGGAAGGATTATATAATGCTACCAATTCAGTGAAGATTATGAATTT 1140  
1081 GTTACCATCGTAGGAAGGATTATATAATGCTACCAATTCAGTGAAGATTATGAATTT 1140  
1141 GACCAATATGAAGAGTTTGTCTGTGTCAGGTTTAAAGCCCTACCCACAATCAGATGT 1200  
1141 GACCAATATGAAGAGTTTGTCTGTGTCAGGTTTAAAGCCCTACCCACAATCAGATGT 1200  
1201 AGGTGGACCTCTCTCGAATAATCATTTCTGTGTCAGCAAAAGGTCCTGTATACGATAC 1260  
1201 AGGTGGACCTCTCTCGAATAATCATTTCTGTGTCAGCAAAAGGTCCTGTATACGATAC 1260  
1261 AGCATATCAAGTTTGTCAATCATAAAGCACCAGCCAGGAGAAATATATTCCTCATGAGAA 1320  
1261 AGCATATCAAGTTTGTCAATCATAAAGCACCAGCCAGGAGAAATATATTCCTCATGAGAA 1320  
1321 AATGATGATGCCCAATTTACCAAAATGTTTCACGCTGAATATAAGAAAGGAAACCTCAAGTG 1380  
1321 AATGATGATGCCCAATTTACCAAAATGTTTCACGCTGAATATAAGAAAGGAAACCTCAAGTG 1380  
1381 CTCGCAGAGCATCGCAAGTACAGGCGTCTGTCTCGGATCGATACCCATTACCATCT 1440  
1381 CTCGCAGAGCATCGCAAGTACAGGCGTCTGTCTCGGATCGATACCCATTACCATCT 1440  
1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCACTGCACAGAGATCACAGAGGA 1500  
1441 TGGACCTGGAAGAGTGTTCAGACAAGTCTCCCACTGCACAGAGATCACAGAGGA 1500  
1501 GTCTGGAATAGAAAGGCTAACGAAAGTGTGTTGGACAGTGGGTGTCGACAGTACTCTA 1560  
1501 GTCTGGAATAGAAAGGCTAACGAAAGTGTGTTGGACAGTGGGTGTCGACAGTACTCTA 1560  
1561 ACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACAAATTCCTTGGC 1620  
1561 ACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTCATACAAATTCCTTGGC 1620  
1621 ACATCTGTGAGACGATCTTTTAAACTCTCCAGGCCCTTCCCTTTCATCCAAAGCAAC 1680  
1621 ACATCTGTGAGACGATCTTTTAAACTCTCCAGGCCCTTCCCTTTCATCCAAAGCAAC 1680  
1681 ATCTCATCTATGCAACAATTTGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTA 1740  
1681 ATCTCATCTATGCAACAATTTGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTA 1740  
1741 ATTTGTCAGAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
1741 ATTTGTCAGAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
1801 ACCGGCTCCTCAGATAATCAGTACTTCTAGCTTGATTTTCAGAGAAATATGAATATGATCTC 1860  
1801 ACCGGCTCCTCAGATAATCAGTACTTCTAGCTTGATTTTCAGAGAAATATGAATATGATCTC 1860  
1861 AAATGGAGTTTCCAGAGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
1861 AAATGGAGTTTCCAGAGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTT 1920  
1921 GGAAGTATGATGAAGCCACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
1921 GGAAGTATGATGAAGCCACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
1981 GCCGTCAAATGCTGAAGAAAAGCAGACAGCTCTGAAAGAGAGGCATCATGTGTCAGAA 2040  
1981 GCCGTCAAATGCTGAAGAAAAGCAGACAGCTCTGAAAGAGAGGCATCATGTGTCAGAA 2040  
2041 CTCAGAGTATGATGACCCAGCTGGGAAGCCACAGAGATTTGTGAACCTGCTGGGGCGTGC 2100  
2041 CTCAGAGTATGATGACCCAGCTGGGAAGCCACAGAGATTTGTGAACCTGCTGGGGCGTGC 2100  
2101 ACACGTCTAGGACCAATTTACTTGATTTTGAATCTGTTGATCTGTTGATCTCTCTCAAC 2160  
2101 ACACGTCTAGGACCAATTTACTTGATTTTGAATCTGTTGATCTGTTGATCTCTCTCAAC 2160

2161 TATCTAAGAAAGTAAAGAGAGAAAATTTTACAGGACTTTGGACAGAGATTTTCAAGGAACAC 2220  
2161 TATCTAAGAAAGTAAAGAGAGAAAATTTTACAGGACTTTGGACAGAGATTTTCAAGGAACAC 2220  
2221 AATTTCAAGTTTACCCACTTTCCAAATCATACATCCAAATTCACAGCATGCTGTTCAGA 2280  
2221 AATTTCAAGTTTACCCACTTTCCAAATCATACATCCAAATTCACAGCATGCTGTTCAGA 2280  
2281 GAAGTTTCAGATACACCCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340  
2281 GAAGTTTCAGATACACCCGGACTCGGATCAATCTCAGGGCTTCATGGGAATTCATTTCAC 2340  
2341 TCTGAAGATCAAAATTTGAATATGAATAACCAAAAGGCTGGAAGAGAGGAGACTTTGAAT 2400  
2341 TCTGAAGATCAAAATTTGAATATGAATAACCAAAAGGCTGGAAGAGAGGAGACTTTGAAT 2400  
2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTGCTATATCAAGTTGCCAAAGGAATGGAATTT 2460  
2401 GTGCTTACATTTGAAGATCTTCTTTGCTTTGCTATATCAAGTTGCCAAAGGAATGGAATTT 2460  
2461 CTGGAATTTAAGTCTGCTGTTTCACAGAGACTGCGCCGCGCAGGAACTGCTTGTCAACCCAC 2520  
2461 CTGGAATTTAAGTCTGCTGTTTCACAGAGACTGCGCCGCGCAGGAACTGCTTGTCAACCCAC 2520  
2521 GGGAAAGTGTGAAGATATGTGACTTTGGATTGCTCGAGATATCATGAGTATTCGAAC 2580  
2521 GGGAAAGTGTGAAGATATGTGACTTTGGATTGCTCGAGATATCATGAGTATTCGAAC 2580  
2581 TATGTTCTCAGGGGCAATGCCCGTCTGCTCTGCTTAAATGGATGGCCCGGAAAGGCTGTTT 2640  
2581 TATGTTCTCAGGGGCAATGCCCGTCTGCTCTGCTTAAATGGATGGCCCGGAAAGGCTGTTT 2640  
2641 GAAGGATCTACACCATTAAGAGTGTCTGTCATATGGAATATTTACTGTGGAAATC 2700  
2641 GAAGGATCTACACCATTAAGAGTGTCTGTCATATGGAATATTTACTGTGGAAATC 2700  
2701 TTCTCACTTGTGTGAATCTTACCTGGGATTCCTGCTGATGCTTAACCTCTCAAAAGCTG 2760  
2701 TTCTCACTTGTGTGAATCTTACCTGGGATTCCTGCTGATGCTTAACCTCTCAAAAGCTG 2760  
2761 ATTCAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATATA 2820  
2761 ATTCAAATGGATTTAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATATA 2820  
2821 ATGCAATCTGCTGGGCTTTTGAATCAAGGAAAGCGCCATCTTCCCTTAATTTGACTTCG 2880  
2821 ATGCAATCTGCTGGGCTTTTGAATCAAGGAAAGCGCCATCTTCCCTTAATTTGACTTCG 2880  
2881 TTTTATAGGATGTCAGCTGGCAGATGCAAGAAAGCGATGATATCAGATGTTGATGGCCGT 2940  
2881 TTTTATAGGATGTCAGCTGGCAGATGCAAGAAAGCGATGATATCAGATGTTGATGGCCGT 2940  
2941 GTTTCGGAATGTCCTCACACCTACCAAAACAGGCGACCTTTTACAGAGAGATGGATTG 3000  
2941 GTTTCGGAATGTCCTCACACCTACCAAAACAGGCGACCTTTTACAGAGAGATGGATTG 3000  
3001 GGGCTACTCTCTCGGAGGCTCAGGTGCAAGTTTGTAGAGGAACAATTTAGTTTAAAG 3060  
3001 GGGCTACTCTCTCGGAGGCTCAGGTGCAAGTTTGTAGAGGAACAATTTAGTTTAAAG 3060  
3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTCAT 3120  
3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTCAT 3120  
3121 CACTAAAAGAAATCTATTTATCAACTGCTTCCACAGACTTTTCTCTAGAAGCCGTCT 3180  
3121 CACTAAAAGAAATCTATTTATCAACTGCTTCCACAGACTTTTCTCTAGAAGCCGTCT 3180  
3181 GCCTTTACTCTGTTTCAAGGGGACTTTTGTAAAATCAATCATCTCTGTCAAGGCAG 3240  
3181 GCCTTTACTCTGTTTCAAGGGGACTTTTGTAAAATCAATCATCTCTGTCAAGGCAG 3240  
3241 GAGGAGCTGATATGAATTTATTTGAGGACTTTGATCTGCAATCCAAAGGCCCTTCTCAGGCCG 3300



Db 1081 ATTGACCAATATGAAGAGTTTTTGTCTGTCTAGGTTTAAAGCCTACCCACAATCAGA 1140  
Qy 1198 TGTACGTGGACCTTCTCTCGAATAATCATTTCTTGTGACCAAAAGGGTCTTGTATAACGGA 1257  
Db 1141 TGTACGTGGACCTTCTCTCGAATAATCATTTCTTGTGACCAAAAGGGTCTTGTATAACGGA 1200  
Qy 1258 TACAGCATATCCAAAGTTTTTGCATATCAATGAAGACCAAGCCAGGAGATATATATCCATGCA 1317  
Db 1201 TACAGCATATCCAAAGTTTTTGCATATCAATGAAGACCAAGCCAGGAGATATATATCCATGCA 1260  
Qy 1318 GAAATGATGATGCCCAATTTACAAAATGTTACGCTGATATATAAGAGAAACCTCAA 1377  
Db 1261 GAAATGATGATGCCCAATTTACAAAATGTTACGCTGATATATAAGAGAAACCTCAA 1320  
Qy 1378 GTGCTCGAGAAGCATCGGCAAGTCAGCGCTCTCTTCTCGGATGGATACCCATTACCA 1437  
Db 1321 GTCTCGCAGAGCTTCGGCAAGTCAGCGCTCTCTTCTCGGATGGATACCCATTACCA 1380  
Qy 1438 TCTTGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATACACAGAA 1497  
Db 1381 TCTTGGACCTGGAAGAAGTGTTCAGACAAGTCTCCCAACTGCACAGAAGAGATACACAGAA 1440  
Qy 1498 GGAGTCTGGAATAGAAAGGCTAACAGAAAGTGTTCGACAGTGGTGTGACAGAGTACT 1557  
Db 1441 GGAGTCTGGAATAGAAAGGCTAACAGAAAGTGTTCGACAGTGGTGTGACAGAGTACT 1500  
Qy 1558 CTAACATGATGTAAGGCATATAAAGGTTCTCTGTCAAGTCTGTGCATACAAATCCCTT 1617  
Db 1501 CTAACATGATGTAAGGCTATAAAGGTTCTCTGTCAAGTCTGTGCATACAAATCCCTT 1560  
Qy 1618 GGCACATCTTGTGAGAGATCTCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGAC 1677  
Db 1561 GGCACATCTTGTGAGAGATCTCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAAGAC 1620  
Qy 1678 AACATCTCATCTATGCAACAATGGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1737  
Db 1621 AACATCTCATCTATGCAACAATGGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680  
Qy 1738 CTAATTTGTCAACAAGTACAAAAGCAATTTAGTGTATGAAAGCCAGCTACAGATGGTACAG 1797  
Db 1681 CTAATTTGTCAACAAGTACAAAAGCAATTTAGTGTATGAAAGCCAGCTACAGATGGTACAG 1740  
Qy 1798 GTGACCGGCTCTCAGATATAGTACTCTCTAGCTTGTGATTTACAGAGAATATGAATATGAT 1857  
Db 1741 GTGACCGGATCTCAGATTTAGTACTCTCTAGCTTGTGATTTACAGAGAATATGAATATGAT 1800  
Qy 1858 CTCAAATGGAGTTTCAAGAGAAAATTTAGATTTGGGAGGTACTAGGATCAGGTGCT 1917  
Db 1801 GTCAATGGGAGTTTCAAGAGAAAATTTAGATTTGGGAGGTACTAGGATCAGGTGCT 1860  
Qy 1918 TTTGGAAAAGTGTATGAAGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG 1977  
Db 1861 TTTGGAAAAGTGTATGAAGCAACAGCTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG 1920  
Qy 1978 GTTGGCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGCAGCTCATGTCA 2037  
Db 1921 GTTACCGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGCAGCTCATGTCA 1980  
Qy 2038 GAACATCAAGATGATGCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGCG 2097  
Db 1981 GAACATCAAGATGATGCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGCG 2040  
Qy 2098 TGCACACTCTCAGGACCAATTTACTTGAATTTTGAATGCTGCTATGCTGATCTCTC 2157  
Db 2041 TGCACACTCTCAGGACCAATTTACTTGAATTTTGAATGCTGCTATGCTGATCTCTC 2100  
Qy 2158 AACTATCTAAGAGTAAAAGAGAAAATTTACAGGACTTTGGACAGAGATTTTCAAGGAA 2217  
Db 2101 AACTATCTAAGAGTAAAAGAGAAAATTTACAGGACTTTGGACAGAGATTTTCAAGGAA 2160  
Qy 2218 CACAATTTTCAAGTTTACCCCACTTTTCCCAATCATCTCCCAATTTCCAGCATGCTGGTTCA 2277  
Db 2161 CACAATTTTCAAGTTTACCCCACTTTTCCCAATCATCTCCCAATTTCCAGCATGCTGGTTCA 2220

Qy 2278 AGAAGATTTCAGATACACCGGACTCGGATCAAAATCTCAGGCTTCATGGGAATTCATTT 2337  
Db 2221 AGAAGATTTCAGATACACCGGACTCGGATCAAAATCTCAGGCTTCATGGGAATTCATTT 2280  
Qy 2338 CACTCTGAAGATGAAATTTGAATATGAAACCAAAAAGGCTGGAAGAAGAGGAGGACTTGG 2397  
Db 2281 CACTCTGAAGATGAAATTTGAATATGAAACCAAAAAGGCTGGAAGAAGAGGAGGACTTGG 2340  
Qy 2398 AATGTGCTTACATTTGAAGATCTTCTTGTGCTTGTGATATCAAGTTGCTCCCAAGAAATGGA 2457  
Db 2341 AATGTGCTTACATTTGAAGATCTTCTTGTGCTTGTGATATCAAGTTGCTCCCAAGAAATGGA 2400  
Qy 2458 TTTCTGGAAATTTAAGTCTGTTCACAGAGACCTGGCCGCCAGGAACCTGCTGTGCACC 2517  
Db 2401 TTTCTGGAAATTTAAGTCTGTTCACAGAGACCTGGCCGCCAGGAACCTGCTGTGCACC 2460  
Qy 2518 CACGGAAGTGGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATCATAGTATGCTC 2577  
Db 2461 CACGGAAGTGGTGAAGATATGTGACTTTTGGATTGGCTCGAGATATCATAGTATGCTC 2520  
Qy 2578 AACTATGTTGTGAGGGCAATGCCGCTGTCTGTAAATGATGGCCGCCGAAAGCCCTG 2637  
Db 2521 AACTATGTTGTGAGGGCAATGCCGCTGTCTGTAAATGATGGCCGCCGAAAGCCCTG 2580  
Qy 2638 TTTGAAGCATCTACACCATTAAGAGTGTCTGTCTATATGGAATATTTACTGTGGGA 2697  
Db 2581 TTTGAAGCATCTACACCATTAAGAGTGTCTGTCTATATGGAATATTTACTGTGGGA 2640  
Qy 2698 ATCTTCTCACTTGGTGTGAATCCTTACCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 2757  
Db 2641 ATCTTCTCACTTGGTGTGAATCCTTACCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 2700  
Qy 2758 CTGATTTCAAAATGGAATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2817  
Db 2701 CTGATTTCAAAATGGAATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760  
Qy 2818 ATAATGCAATCCTGCTGGGCTTTTGAATCAAGGAACGGCCATCTTCCCTAATTTGACT 2877  
Db 2761 ATAATGCAATCCTGCTGGGCTTTTGAATCAAGGAACGGCCATCTTCCCTAATTTGACT 2820  
Qy 2878 TCGTTTTTAGGATGTGATGCTGAGATGCAAGAAGGAGTATATCAGAAATGATGATGATGATGATGAT 2937  
Db 2821 TCGTTTTTAGGATGTGATGCTGAGATGCAAGAAGGAGTATATCAGAAATGATGATGATGATGATGAT 2880  
Qy 2938 CGTGTTCGGAATGTCTCAGACCTTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGAT 2997  
Db 2881 CGTGTTCGGAATGTCTCAGACCTTACCAAAACAGGCGACCTTTTCAGCAGAGAGATGAT 2940  
Qy 2998 TTGGGGCTACTCTCTCCGAGGCTCAGTCAAGATTCGTAGAGGAACAATTTAGTTTGA 3057  
Db 2941 TTGGGGCTACTCTCTCCGAGGCTCAGTCAAGATTCGTAGAGGAACAATTTAGTTTGA 3000  
Qy 3058 AGGACTTATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAGATTAATTT 3117  
Db 3001 AGGACTTATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAGGTTAAATTT 3060  
Qy 3118 CATCACTAAAAGAAAATCTATTATCACTGTGCTTCCAGAGACTTTTCTCTAGAGCGG 3177  
Db 3061 CATCACTAAAAGAAAATCTATTATCACTGTGCTTCCAGAGACTTTTCTCTAGAGCGG 3120

## RESULT 14

US-08-446-648-22  
; Sequence 22, Application US/08446648  
; Patent No. 6331302  
; GENERAL INFORMATION:  
; APPLICANT: Genentech, Inc.  
; APPLICANT: Bennett, Brian D.  
; APPLICANT: Goeddel, David  
; APPLICANT: Lee, James M.  
; APPLICANT: Matthews, William  
; APPLICANT: Tsai, Siao Ping



QY 1618 GGCACATCTTGTGAGAGCATCCCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGAC 1677  
DB 1561 GGCACATCTTGTGAGAGCATCCCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCAAGAC 1620  
QY 1678 AACATCTCATCTATCAACAATTTGGTGTGTCTCTCTCTCATTTGCTGTTTAAACCTG 1737  
DB 1621 AACATCTCATCTATCAACAATTTGGTGTGTCTCTCTCTCATTTGCTGTTTAAACCTG 1680  
QY 1738 CTAATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG 1797  
DB 1681 CTAATTTGTCAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAG 1740  
QY 1798 GTGACGGCTCTCAGATAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGAT 1857  
DB 1741 GTGACGGGATCTCAGATTAATGAGTACTTCTACGTTGATTTTCAGAGAATATGAATATGAT 1800  
QY 1858 CTCAAATGGAGTTTCCACAGAAAATTTAGATTTTGGGAAGGTACTAGGATCAGTGTCT 1917  
DB 1801 GTCAAATGGAGTTTCCACAGAAAATTTAGATTTTGGGAAGGTACTAGGATCAGTGTCT 1860  
QY 1918 TTTGGAAGTGTGATGACGCAACAGCTTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG 1977  
DB 1861 TTTGGAAGTGTGATGACGCAACAGCTTTATGGAATTAGCAAAACAGGAGTCTCAATCCAG 1920  
QY 1978 GTTGGCGTCAAAATGCTGAAAGAAAAGACAGACAGCTCTGAAAGAGAGGCACTCATCTCA 2037  
DB 1921 GTTACCGTCAAAATGCTGAAAGAAAAGACAGACAGCTCTGAAAGAGAGGCACTCATCTCA 1980  
QY 2038 GAACCTCAAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGCG 2097  
DB 1981 GAACCTCAAGATGATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGCG 2040  
QY 2098 TGCACACTGTGAGGACCAATTTACTTGTATTTTGAATTTGTGATGTTGCTATGCTGATCTTCTC 2157  
DB 2041 TGCACACTGTGAGGACCAATTTACTTGTATTTTGAATTTGTGATGTTGCTATGCTGATCTTCTC 2100  
QY 2158 AACTATCTAAGAGTAAAGAGAAAAATTTACAGGACTTTGGACAGAGATTTTCAAGGAA 2217  
DB 2101 AACTATCTAAGAGTAAAGAGAAAAATTTACAGGACTTTGGACAGAGATTTTCAAGGAA 2160  
QY 2218 CACAATTTTCAGTTTACCCCACTTTCCATCATCATCCCAATTTCCAGCATGCTCGTTTCA 2277  
DB 2161 CACAATTTTCAGTTTACCCCACTTTCCATCATCATCCCAATTTCCAGCATGCTCGTTTCA 2220  
QY 2278 AGAAGATTCAGATACACCCGGACTCGGATCAATCTCAGGCTTCATGGGAATTCATTT 2337  
DB 2221 AGAAGATTCAGATACACCCGGACTCGGATCAATCTCAGGCTTCATGGGAATTCATTT 2280  
QY 2338 CACTCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAGAGGAGGACTTG 2397  
DB 2281 CACTCTGAAGATGAAATTTGAATATGAAACCAAAAGGCTGGAAGAGAGGAGGACTTG 2340  
QY 2398 AATGTGCTTACATTTGAAGATCTTCTTGTGTCATATCAAGTTGCCAAAGGAATGGAA 2457  
DB 2341 AATGTGCTTACATTTGAAGATCTTCTTGTGTCATATCAAGTTGCCAAAGGAATGGAA 2400  
QY 2458 TTTCTGGAATTTAAGTCGTGTTCACAGAGACTGGCCGCCAGGAGGCTGTTCACC 2517  
DB 2401 TTTCTGGAATTTAAGTCGTGTTCACAGAGACTGGCCGCCAGGAGGCTGTTCACC 2460  
QY 2518 CACGGGAAAGTGGTGAAGATATGTACTTTGGATTGGCTCGAGATATCATGAGTGTATCC 2577  
DB 2461 CACGGGAAAGTGGTGAAGATATGTACTTTGGATTGGCTCGAGATATCATGAGTGTATCC 2520  
QY 2578 AACTATGTTGTGAGGGGCAATGCCCTGTCCCTGTAAATGGATGGCCGCCAAAGCCTG 2637  
DB 2521 AACTATGTTGTGAGGGGCAATGCCCTGTCCCTGTAAATGGATGGCCGCCAAAGCCTG 2580  
QY 2638 TTTGAAGGATCTACACCAATTAAGAGTGTGCTGTGTCATATGATATATCTGTGGGA 2697  
DB 2581 TTTGAAGGATCTACACCAATTAAGAGTGTGCTGTGTCATATGATATATCTGTGGGA 2640  
QY 2698 ATCTTCTACTTGGTGTGAATCTTACCCTGGCATTCCGCTTGTGATCTTAACCTTCAAA 2757

DB 2641 ATCTTCTACATTTGTTGTAATCCCTTACCTTGGCATTCGGTGTGATGCTTAACCTTCTACAA 2700  
QY 2758 CTGATTTCAAAATGATTTAAATGATCAGCCATTTTATGTCTACAGAAAGAAATATACATT 2817  
DB 2701 CTGATTTCAAAATGATTTAAATGATCAGCCATTTTATGTCTACAGAAAGAAATATACATT 2760  
QY 2818 ATAATGAATCTCTCTGGCTTTTGTGACTCAAGGAAACGGCCATCTTCCCTAAATTTGACT 2877  
DB 2761 ATAATGAATCTCTCTGGCTTTTGTGACTCAAGGAAACGGCCATCTTCCCTAAATTTGACT 2820  
QY 2878 TCGTTTTTAGGATGTCTGAGCTGGCAGATGCAGAAAGAGGATGTATCAGAATCTGGATGGC 2937  
DB 2821 TCGTTTTTAGGATGTCTGAGCTGGCAGATGCAGAAAGAGGATGTATCAGAATCTGGATGGC 2880  
QY 2938 CGTGTTCGGAATGTCTCTACACCTTACCAAAACAGGCAACCTTTCAGCAGAGATGGAT 2997  
DB 2881 CGTGTTCGGAATGTCTCTACACCTTACCAAAACAGGCAACCTTTCAGCAGAGATGGAT 2940  
QY 2998 TTGGGGTACTCTCTCGCAGGCTCAGTCCGAAGATTCGTAGAGAAACAATTTAGTTTTA 3057  
DB 2941 TTGGGGTACTCTCTCGCAGGCTCAGTCCGAAGATTCGTAGAGAAACAATTTAGTTTTA 3000  
QY 3058 AGGACTTATCTCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAAGATTAATTT 3117  
DB 3001 AGGACTTATCTCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAAACAAGTAAATTT 3060  
QY 3118 CATCACTAAAAGAAAATCTATTATCAACTGCTGTCTTACCAGACTTTTCTCTAGAGCGC 3177  
DB 3061 CATCACTAAAAGAAAATCTATTATCAACTGCTGTCTTACCAGACTTTTCTCTAGAGCGC 3120

## RESULT 15

PCT-US95-04228-22  
: Sequence 22, Application PC/TUS9504228  
: GENERAL INFORMATION:  
: APPLICANT: Genentech, Inc.  
: APPLICANT: Bennett, Brian D.  
: APPLICANT: Goeddel, David  
: APPLICANT: Lee, James M.  
: APPLICANT: Matthews, William  
: APPLICANT: Tsai, Siao Ping  
: APPLICANT: Wood, William I.  
: TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES  
: NUMBER OF SEQUENCES: 45  
: CORRESPONDENCE ADDRESS:  
: STREET: 460 Point San Bruno Blvd  
: CITY: South San Francisco  
: STATE: California  
: COUNTRY: USA  
: ZIP: 94080  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: 5.25 inch, 360 kb floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: patin (Genentech)  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: PCT/US95/04228  
: FILING DATE:  
: CLASSIFICATION:  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 08/22616  
: FILING DATE: 04-APR-1994  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Wendy M. Lee  
: REGISTRATION NUMBER: 00,000  
: REFERENCE/DOCKET NUMBER: 821P3PCT  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 415/225-1994  
: TELEFAX: 415/952-9881  
: TELEX: 910/371-7168  
: INFORMATION FOR SEQ ID NO: 22:



Qy	2098	TGCACACTGTCAGAGACCAATTTACTTGATTTTGGAACTACTGTTGCTATGCTGATGCTCTC	2155
Db	2041	TGCACACTGTCAGAGACCAATTTACTTGATTTTGGAACTACTGTTGCTATGCTGATGCTCTC	2100
Qy	2158	AACATATCTAAGAAGCTAAAGAGAAAAATTTACAGAGGACTTGGACAGAGATTTTCAAGGAA	2217
Db	2101	AACATCTAAGAAGCTAAAGAGAAAAATTTACAGAGGACTTGGACAGAGATTTTCAAGGAA	2160
Qy	2218	CACAATTTACGTTTTTATCCCCACATTTCCAAATCACATCCAAATTTCCAGCATGCTGTTCA	2277
Db	2161	CACAATTTACGTTTTTATCCCCACATTTCCAAATCACATCCAAATTTCCAGCATGCTGTTCA	2220
Qy	2278	AGAGAAGTTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGAAATTCATTT	2337
Db	2221	AGAGAAGTTTCAGATACACCCGGACTCGGATCAAAATCTCAGGGCTTCATGGAAATTCATTT	2280
Qy	2338	CACCTCAAGATGAAATGTAATATGAAAAACCAAAAAGGCTGGAAGAGAGGAGGACTTG	2397
Db	2281	CACCTCAAGATGAAATGTAATATGAAAAACCAAAAAGGCTGGAAGAGAGGAGGACTTG	2340
Qy	2398	AATGTGCTTACATTTGAAGATCTCTTTGCTTGTGCATATCAAGTTGCGCAAGGAATGGAA	2457
Db	2341	AATGTGCTTACATTTGAAGATCTCTTTGCTTGTGCATATCAAGTTGCGCAAGGAATGGAA	2400
Qy	2458	TTTCTGGAATTTAAGTCGTGTGTTCACAGAGACCTGGCCGCCAGAACGCTGCTGTGCACC	2517
Db	2401	TTTCTGGAATTTAAGTCGTGTGTTCACAGAGACCTGGCCGCCAGAACGCTGCTGTGCACC	2460
Qy	2518	CACGGGAAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATAGTGAATCC	2577
Db	2461	CACGGGAAAAGTGGTGAAGATATGTGACTTTGGATTTGGCTCGAGATATCATAGTGAATCC	2520
Qy	2578	AACATATGCTGTCAAGGGCAATGCCGCTCCCTGTAAATGGATGGCCCCGAAAAGCCCTG	2637
Db	2521	AACATATGCTGTCAAGGGCAATGCCGCTCCCTGTAAATGGATGGCCCCGAAAAGCCCTG	2580
Qy	2638	TTTGAAGGCATCTACACCATTAAGATGATGTCTGTGTCATATGGAATATTACTGTGGGAA	2697
Db	2581	TTTGAAGGCATCTACACCATTAAGATGATGTCTGTGTCATATGGAATATTACTGTGGGAA	2640
Qy	2698	ATCTTCTCACTTGGTGTGAATCCTTACCCTGGCAATCCCGTTGATGCTCACTTCTCTACAAA	2757
Db	2641	ATCTTCTCACTTGGTGTGAATCCTTACCCTGGCAATCCCGTTGATGCTCACTTCTCTACAAA	2700
Qy	2758	CTGATTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTACACAAGAAATATACATT	2817
Db	2701	CTGATTCAAAATGGATTTAAATGGATCAGCCATTTTATGCTACACAAGAAATATACATT	2760
Qy	2818	ATAATGCAATCCCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCCCTTAATTTGACT	2877
Db	2761	ATAATGCAATCCCTGCTGGGCTTTTGACTCAAGGAAACGGCCATCCCTTAATTTGACT	2820
Qy	2878	TCGTTTTTAGATGTCAGCTGGCAGATGCAGAGAAGCGATGTATCAGAAATGTGGATGGC	2937
Db	2821	TCGTTTTTAGATGTCAGCTGGCAGATGCAGAGAAGCGATGTATCAGAAATGTGGATGGC	2880
Qy	2938	CGTGTTCGGAATGTCCTCACACCTACCAAAACAGCGGACCTTTCAGCAGAGAGATGGAT	2997
Db	2881	CGTGTTCGGAATGTCCTCACACCTACCAAAACAGCGGACCTTTCAGCAGAGAGATGGAT	2940
Qy	2998	TTGGGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTTTA	3057
Db	2941	TTGGGGCTACTCTCTCCGAGGCTCAGGTCGAAGATTCGTAGAGGAACAATTTAGTTTTTA	3000
Qy	3058	AGGACTTATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTT	3117
Db	3001	AGGACTTATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAGATTAATTT	3060
Qy	3118	CATCACTAAAGAAAAATCTATTATCAACTGCTGCTTCCACGACTTTTCTCTAGAGCGG	3177
Db	3061	CATCACTAAAGAAAAATCTATTATCAACTGCTGCTTCCACGACTTTTCTCTAGAGCGG	3120



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 11:51:23 ; Search time 403.768 seconds  
(without alignments)  
11449.495 Million cell updates/sec

Title: US-09-919-408-3

Perfect score: 3501

Sequence: 1 CGAGGCGCATCCGAGGCT.....AAAAAAAAAAAAAAAAAAAA 3501.

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3501	100.0	3501	10	US-09-919-408-3
2	3501	100.0	3501	10	US-09-872-136-3
3	3083.2	88.1	3120	10	US-09-982-610-22
4	2227.2	63.6	3453	10	US-09-919-408-1
5	2227.2	63.6	3453	10	US-09-872-136-1
6	633	18.1	645	9	US-09-796-692-7538
7	633	18.1	645	9	US-10-040-862-7538
8	608.4	17.4	610	9	US-09-796-692-8085
9	608.4	17.4	610	9	US-10-040-862-8085
10	606.4	17.3	608	9	US-09-796-692-8132
11	606.4	17.3	608	9	US-10-040-862-8132
12	605	17.3	605	9	US-09-796-692-8927
13	605	17.3	605	9	US-10-040-862-8927
14	598.4	17.1	601	9	US-09-796-692-8273
15	598.4	17.1	601	9	US-10-040-862-8273
16	597	17.1	597	9	US-09-796-692-7976
17	597	17.1	597	9	US-10-040-862-7976
18	595.4	17.0	597	9	US-09-796-692-9010
19	595.4	17.0	597	9	US-10-040-862-9010

20	593.2	16.9	598	9	US-09-796-692-8566	Sequence 8566, Ap
21	593.2	16.9	598	9	US-10-040-862-8566	Sequence 8566, Ap
22	590.8	16.9	594	9	US-09-796-692-8135	Sequence 8135, Ap
23	590.8	16.9	594	9	US-10-040-862-8135	Sequence 8135, Ap
24	590.4	16.9	605	9	US-09-796-692-8047	Sequence 8047, Ap
25	590.4	16.9	605	9	US-10-040-862-8047	Sequence 8047, Ap
26	583.8	16.7	589	9	US-09-796-692-8630	Sequence 8630, Ap
27	583.8	16.7	589	9	US-10-040-862-8630	Sequence 8630, Ap
28	582.4	16.1	564	9	US-09-796-692-9025	Sequence 9025, Ap
29	582.4	16.1	564	9	US-10-040-862-9025	Sequence 9025, Ap
30	554.4	15.8	556	9	US-09-796-692-7883	Sequence 7883, Ap
31	554.4	15.8	556	9	US-10-040-862-7883	Sequence 7883, Ap
32	541	15.5	550	9	US-09-796-692-8494	Sequence 8494, Ap
33	541	15.5	550	9	US-10-040-862-8494	Sequence 8494, Ap
34	539	15.4	539	9	US-09-796-692-8255	Sequence 8255, Ap
35	539	15.4	539	9	US-10-040-862-8255	Sequence 8255, Ap
C	510	14.6	510	9	US-09-796-692-8268	Sequence 8268, Ap
36	510	14.6	510	9	US-10-040-862-8268	Sequence 8268, Ap
37	510	14.6	510	9	US-09-796-692-8032	Sequence 8032, Ap
38	502.8	14.4	508	9	US-10-040-862-8032	Sequence 8032, Ap
39	502.8	14.4	508	9	US-09-796-692-8408	Sequence 8408, Ap
40	491.4	14.0	497	9	US-10-040-862-8408	Sequence 8408, Ap
41	491.4	14.0	497	9	US-09-796-692-8114	Sequence 8114, Ap
42	478.4	13.7	606	9	US-10-040-862-8114	Sequence 8114, Ap
43	478.4	13.7	606	9	US-09-796-692-9031	Sequence 9031, Ap
44	446.4	12.8	581	9	US-10-040-862-9031	Sequence 9031, Ap
45	446.4	12.8	581	9	US-09-796-692-9031	Sequence 9031, Ap

#### ALIGNMENTS

#### RESULT 1

US-09-919-408-3

; Sequence 3, Application US/09919408

; Patent No. US2002007207A1

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor R.

; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL RECEPTORS AND THEIR LIGANDS

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ImClone Systems Incorporated

; STREET: 180 Varick Street

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10014

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/919,408

; FILING DATE: 31-Jul-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/977,451

; FILING DATE: <Unknown>

; APPLICATION NUMBER: US 07/906,397

; FILING DATE: 26-JUN-1992

; APPLICATION NUMBER: US PCT/US92/05401

; FILING DATE: 26-JUN-1992

; APPLICATION NUMBER: TW 81102961

; FILING DATE: 15-APR-1992

; APPLICATION NUMBER: US PCT/US92/02750

; FILING DATE: 02-APR-1992

; APPLICATION NUMBER: US 07/813,593

; FILING DATE: 24-DEC-1991

; APPLICATION NUMBER: US 07/793,065

; FILING DATE: 15-NOV-1991

; APPLICATION NUMBER: US 07/728,913

; FILING DATE: 28-JUN-1991

APPLICATION NUMBER: US 07/679,666

FILING DATE: 02-APR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Feit, Irving N.

REGISTRATION NUMBER: 28 601

REFERENCE/DOCKET NUMBER: LEM-3-7P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-645-1405

TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 3501 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

FEATURE:

NAME/KEY: CDS

LOCATION: 58...3039

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: 139...3036

FEATURE:

NAME/KEY: sig\_peptide

LOCATION: 58...138

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

us-09-919-408-3

Query Match 100.0%; Score 3501; DB 10; Length 3501;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CGAGCGGCATCCGAGGCTGGCGCGCCCTCGGGGACCCCGGCTCCGGAGGCCATG	60
DB	1	CGAGCGGCATCCGAGGCTGGCGCGCCCTCGGGGACCCCGGCTCCGGAGGCCATG	60
QY	61	CCGGCGTGGCGCGAGCGCGGACCGGCGCGCTGCTGTTGTTTCTGCAATGATA	120
DB	61	CCGGCGTGGCGCGAGCGCGGACCGGCGCGCTGCTGTTGTTTCTGCAATGATA	120
QY	121	TTTGGGACTATTACAAATCAAGATGCTGCTGATCAAGTGTGTTTAAATCAATCAAG	180
DB	121	TTTGGGACTATTACAAATCAAGATGCTGCTGATCAAGTGTGTTTAAATCAATCAAG	180
QY	181	AACAATGATTCATAGTGGGGAAGTATCATATCCCATGGTATCAGAAATCCCGGAA	240
DB	181	AACAATGATTCATAGTGGGGAAGTATCATATCCCATGGTATCAGAAATCCCGGAA	240
QY	241	GACCTCGGCTGCGTGGAGCCAGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGTG	300
DB	241	GACCTCGGCTGCGTGGAGCCAGAGCTCAGGGACAGTGTACGAAGCTGCCGCTGTG	300
QY	301	GAAGTGGATGATCTGCTTCCATCAACATGCAAGTGTGCTGATGCCCGCAGGGAACAT	360
DB	301	GAAGTGGATGATCTGCTTCCATCAACATGCAAGTGTGCTGATGCCCGCAGGGAACAT	360
QY	361	TCCGTGCTCTGGTCTTTAAGCACAGCTCCCTGAATTTGCGAGCCATTTGATTTACAA	420
DB	361	TCCGTGCTCTGGTCTTTAAGCACAGCTCCCTGAATTTGCGAGCCATTTGATTTACAA	420
QY	421	AACAGAGGAGTGTGTTCCATGCTCATTTTGAATATGACAGAAACCAAGCTGGAGATAC	480
DB	421	AACAGAGGAGTGTGTTCCATGCTCATTTTGAATATGACAGAAACCAAGCTGGAGATAC	480
QY	481	CTACTTTTATTCAGAGTGAAGTACCAATTTACACAATATTTGTACAGTGAATTAAGA	540
DB	481	CTACTTTTATTCAGAGTGAAGTACCAATTTACACAATATTTGTACAGTGAATTAAGA	540
QY	541	AATACCCCTCTTACACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACAGGAGC	600
DB	541	AATACCCCTCTTACACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACAGGAGC	600

DB	541	AATACCCCTCTTACACATTAAGAAGACCTTACTTTAGAAAAATGAAAAACAGGAGC	600
QY	601	CTGGTCTGCATATCTGAGAGCCTTCCAGAGCCGATCGTGAATGGTGTCTTTCGCAATCA	660
DB	601	CTGGTCTGCATATCTGAGAGCCTTCCAGAGCCGATCGTGAATGGTGTCTTTCGCAATCA	660
QY	661	CAGGGGAAGCTGTAAAGAAAGAGTCCAGCTGTTGTTTAAAAAGAGGAAAAAGTGCTT	720
DB	661	CAGGGGAAGCTGTAAAGAAAGAGTCCAGCTGTTGTTTAAAAAGAGGAAAAAGTGCTT	720
QY	721	CATGAATTTATTTGGGACGACATAGTGTCTGTGCCAGAAATGAAGTGGGAGGAAATGC	780
DB	721	CATGAATTTATTTGGGACGACATAGTGTCTGTGCCAGAAATGAAGTGGGAGGAAATGC	780
QY	781	ACCAGGCTGTTCCACATAGATCTAAATCAAACTCTCAGACCACATGTCACACATATTT	840
DB	781	ACCAGGCTGTTCCACATAGATCTAAATCAAACTCTCAGACCACATGTCACACATATTT	840
QY	841	CTTAAAGTAGGGGACCTTATGGATAAGGTGCAAAAGCTGTTTCATGTGAACCATGGATTC	900
DB	841	CTTAAAGTAGGGGACCTTATGGATAAGGTGCAAAAGCTGTTTCATGTGAACCATGGATTC	900
QY	901	GGGCTCACCTGGGAATTAGAAAAACAAGCACTCGAGGAGGCAACTACTTTGAGATGAGT	960
DB	901	GGGCTCACCTGGGAATTAGAAAAACAAGCACTCGAGGAGGCAACTACTTTGAGATGAGT	960
QY	961	ACCTATTCAACAAACAGAACTATGATACGGATTTCTGTTGCTTTTGTATCATCAGTGGCA	1020
DB	961	ACCTATTCAACAAACAGAACTATGATACGGATTTCTGTTGCTTTTGTATCATCAGTGGCA	1020
QY	1021	AGAAAGCACCGGATACACTTGTCTCTTCAAGCATCCAGTCAATCAGCTTTG	1080
DB	1021	AGAAAGCACCGGATACACTTGTCTCTTCAAGCATCCAGTCAATCAGCTTTG	1080
QY	1081	GTTTACCATCGTAGGAAAGGATTTATAATGCTTACCAATTCAGTGAAGATTTAGAAAT	1140
DB	1081	GTTTACCATCGTAGGAAAGGATTTATAATGCTTACCAATTCAGTGAAGATTTAGAAAT	1140
QY	1141	GACCAATTAAGAGATTTGTTTCTGTCAGGTTTAAAGCTTACCCACAAATCAGATGT	1200
DB	1141	GACCAATTAAGAGATTTGTTTCTGTCAGGTTTAAAGCTTACCCACAAATCAGATGT	1200
QY	1201	ACGTGACCTTCTCGAAATCATTTCTTGTGAGCAAAAGGGTCTGTATACGGATAC	1260
DB	1201	ACGTGACCTTCTCGAAATCATTTCTTGTGAGCAAAAGGGTCTGTATACGGATAC	1260
QY	1261	AGCATATCAAGTTTGCATATCATAAGCACAGCAGGAGAAATATATATCCATGCAGAA	1320
DB	1261	AGCATATCAAGTTTGCATATCATAAGCACAGCAGGAGAAATATATATCCATGCAGAA	1320
QY	1321	AATGATGATGCCCAATTTACCAAAATGTTCCAGCTGAATATAAGAGGAAACCTCAAGTG	1380
DB	1321	AATGATGATGCCCAATTTACCAAAATGTTCCAGCTGAATATAAGAGGAAACCTCAAGTG	1380
QY	1381	CTCGAGAGCATCGGCAAGTCCAGCTGCTGTTTTCGGATGGATACCATTTACCATCT	1440
DB	1381	CTCGAGAGCATCGGCAAGTCCAGCTGCTGTTTTCGGATGGATACCATTTACCATCT	1440
QY	1441	TGGACCTGGAAGAGTGTTCAGACAACTCTCCCAACTGCACAGAGAGATCACAGAGGA	1500
DB	1441	TGGACCTGGAAGAGTGTTCAGACAACTCTCCCAACTGCACAGAGAGATCACAGAGGA	1500
QY	1501	GTCTGGAATAGAAGGCTAACAGAAAAGTGTGTCAGTGGGTGTCGAGCAGTACTCTA	1560
DB	1501	GTCTGGAATAGAAGGCTAACAGAAAAGTGTGTCAGTGGGTGTCGAGCAGTACTCTA	1560
QY	1561	AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTGCATCAATTCCTTGGC	1620
DB	1561	AACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGTGCATCAATTCCTTGGC	1620
QY	1621	ACATCTGTGAGAGCATCTTTTAAACTCTCCAGGCCCCCTTCCCTTTCATCCAGACAAC	1680
DB	1621	ACATCTGTGAGAGCATCTTTTAAACTCTCCAGGCCCCCTTCCCTTTCATCCAGACAAC	1680

1681 ATCTCATCTTATCAACAATTGGTGTGTTGTCCTCTTCTATTGTCGTTTTAAACCCCTGCTA 1740  
1681 ATCTCATCTTATCAACAATTGGTGTGTTGTCCTCTTCTATTGTCGTTTTAAACCCCTGCTA 1740  
1741 ATTTGTCACAAGTACAAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800  
1741 ATTTGTCACAAGTACAAAAAGCAATTTAGGTATGAAGCCAGCTACAGATGGTACAGGTG 1800  
1801 ACCGGCTCTCAGATAATGAGTACTTCTAGCTTGATTTTCAGAGAAATATGAATATGATCTC 1860  
1801 ACCGGCTCTCAGATAATGAGTACTTCTAGCTTGATTTTCAGAGAAATATGAATATGATCTC 1860  
1861 AAATGGGAGCTTCCAAAGAGAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT 1920  
1861 AAATGGGAGCTTCCAAAGAGAAATTTAGAGTTTGGGAAGTACTAGGATCAGGTGCTTTT 1920  
1921 GGAAGAGTGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
1921 GGAAGAGTGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGGAGTCTCAATCCAGGTT 1980  
1981 GCCGTCAAAATGCTGAAGAAAGAGAGAGAGTCTGAAAGAGAGGAGTCTGATGTCAGAA 2040  
1981 GCCGTCAAAATGCTGAAGAAAGAGAGAGAGTCTGAAAGAGAGGAGTCTGATGTCAGAA 2040  
2041 CTCAGATGATGACCCAGCTGGGAAGCCAGGAGATTTGTAACCTGCTGGGGGGGCTG 2100  
2041 CTCAGATGATGACCCAGCTGGGAAGCCAGGAGATTTGTAACCTGCTGGGGGGGCTG 2100  
2101 ACACCTGTCAGGACCAATTTACTTGTGTTGATGTTGTTGATGTTGTTGATGTTGTTCAAC 2160  
2101 ACACCTGTCAGGACCAATTTACTTGTGTTGATGTTGTTGATGTTGTTGATGTTGTTCAAC 2160  
2161 TATCTAAGAGTAAAGAGAGAAATTTACAGAGACTTGGACAGAGATTTTCAAGAGAAC 2220  
2161 TATCTAAGAGTAAAGAGAGAAATTTACAGAGACTTGGACAGAGATTTTCAAGAGAAC 2220  
2221 AATTTTCAGTGTTCACCCACTTTCACATCCAAATTTCCAGCATGCTGCTGTTCAAGA 2280  
2221 AATTTTCAGTGTTCACCCACTTTCACATCCAAATTTCCAGCATGCTGCTGTTCAAGA 2280  
2281 GAAGTTCAGATACACCCGACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTTCAC 2340  
2281 GAAGTTCAGATACACCCGACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTTCAC 2340  
2341 TCTGAGATGAATTTGAATATGAAGAAAGCAATTTGATGTTGATGTTGATGTTGATGTTGAT 2400  
2341 TCTGAGATGAATTTGAATATGAAGAAAGCAATTTGATGTTGATGTTGATGTTGATGTTGAT 2400  
2401 GTGCTTACATTTGAAGATCTTCTTCTGTTGATATCAAGTTGCCAAAGGAATGGAATTT 2460  
2401 GTGCTTACATTTGAAGATCTTCTTCTGTTGATATCAAGTTGCCAAAGGAATGGAATTT 2460  
2461 CTGGAATTTAAGTTCGTTGTTTACAGAGACTTGGCCGCCAGGAACTGTTTGCACCCAC 2520  
2461 CTGGAATTTAAGTTCGTTGTTTACAGAGACTTGGCCGCCAGGAACTGTTTGCACCCAC 2520  
2521 GGAAGAGTGGTGAAGATATGAGTGTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 2580  
2521 GGAAGAGTGGTGAAGATATGAGTGTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 2580  
2581 TATGTTGTCAGGGGAATCCCGCTGCTGCTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 2640  
2581 TATGTTGTCAGGGGAATCCCGCTGCTGCTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 2640  
2641 GAAGGATCTACACCATTAAGAGTGTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 2700  
2641 GAAGGATCTACACCATTAAGAGTGTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 2700  
2701 TTCTCAGTGTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 2760  
2701 TTCTCAGTGTGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTGAT 2760

2761 ATTCAAAATGGATTTAAAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATTATA 2820  
2761 ATTCAAAATGGATTTAAAAATGGATCAGCCATTTTATGCTACAGAGAAATATACATTATA 2820  
2821 ATGCAATCTCTGCTGGGCTTTTGTACTCAAGGAAAGCGCATCTTCCCTTAATTTGACTTCG 2880  
2821 ATGCAATCTCTGCTGGGCTTTTGTACTCAAGGAAAGCGCATCTTCCCTTAATTTGACTTCG 2880  
2881 TTTTGTAGGATGTACGTGGGATGTCAGAGAAAGCGGATGTATCAGAAATGTGGATGCCGT 2940  
2881 TTTTGTAGGATGTACGTGGGATGTCAGAGAAAGCGGATGTATCAGAAATGTGGATGCCGT 2940  
2941 GTTTCGAAGTGTCTCTACACCTTACCAAAACAGGCGCTTTTCAGCAGAGATGATTTG 3000  
2941 GTTTCGAAGTGTCTCTACACCTTACCAAAACAGGCGCTTTTCAGCAGAGATGATTTG 3000  
3001 GGGTACTCTCTCCGAGGCTCAGTGCAGATTCGTAGAGAAACAAATTTAGTTTAAAG 3060  
3001 GGGTACTCTCTCCGAGGCTCAGTGCAGATTCGTAGAGAAACAAATTTAGTTTAAAG 3060  
3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAAGATTAATTTTCA 3120  
3061 ACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTACCAAAACAAAGATTAATTTTCA 3120  
3121 CACTAAAGAAATCTATTATCACTGCTTCCAGAGCTTTTCTCTAGAGCCGTCT 3180  
3121 CACTAAAGAAATCTATTATCACTGCTTCCAGAGCTTTTCTCTAGAGCCGTCT 3180  
3181 GCGTTTACTCTGTTTCAAGGAGCTTTTGTAAAATCAAAATCATCTCTCAGAGGCG 3240  
3181 GCGTTTACTCTGTTTCAAGGAGCTTTTGTAAAATCAAAATCATCTCTCAGAGGCG 3240  
3241 GAGAGCTGATTAATGAACCTTTATGGAGCATTTGATGTCATCAAGGCTTCTCAGGCG 3300  
3241 GAGAGCTGATTAATGAACCTTTATGGAGCATTTGATGTCATCAAGGCTTCTCAGGCG 3300  
3301 GCTTGAAGTGTGTCACCTGAGTACAGTATATCTTGTAAATACATAAACAAGC 3360  
3301 GCTTGAAGTGTGTCACCTGAGTACAGTATATCTTGTAAATACATAAACAAGC 3360  
3361 ATTTGCTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATATATGATAA 3420  
3361 ATTTGCTAAGGAGAGCTAATATGATTTTAAAGTCTATGTTTAAATATATGATAA 3420  
3421 TTTTTCAGCTATTTAGTGTATATTTTATGGTGGGAAATAAATTTCTACTACAGAAAA 3480  
3421 TTTTTCAGCTATTTAGTGTATATTTTATGGTGGGAAATAAATTTCTACTACAGAAAA 3480  
3481 AAAAAAAAAAAAAAAAAAAAAA 3501  
3481 AAAAAAAAAAAAAAAAAAAAAA 3501

## RESULT 2

US-09-872-136-3

; Sequence 3, Application US/09872136

; Patent No. US20020119545A1

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor R.

; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

; RECEPTORS AND THEIR LIGANDS

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; STREET: 180 Varick Street

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10014

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS



|||||  
1261 AGCATATCCAAGTTTGCATATATAGCACCAGCCAGGAGATATATATTCATTCAGAGAA 1320  
QY  
1321 AATGATGATGCCAAATTTACCAAAATGTTCCAGCTGAATATATAGAGAAACCTCAAGTG 1380  
Db  
1321 AATGATGATGCCAAATTTACCAAAATGTTCCAGCTGAATATATAGAGAAACCTCAAGTG 1380  
QY  
1381 CTCGCAGAACATCGGCAAGTCTAGGCGTCTCTGTTCTCGGATGGATACCCATTACCATCT 1440  
Db  
1381 CTCGCAGAACATCGGCAAGTCTAGGCGTCTCTGTTCTCGGATGGATACCCATTACCATCT 1440  
QY  
1441 TGGACCTGGAAGAAGTGTTCAGACAAAGTCTCCCAACTGCAAGAGAGATCACAGAAGGA 1500  
Db  
1441 TGGACCTGGAAGAAGTGTTCAGACAAAGTCTCCCAACTGCAAGAGAGATCACAGAAGGA 1500  
QY  
1501 GTCTGGAATAGAAAGGCTACAGAAAGTGTGTTGGACAGTGGGTGTCGAGAGTACTCTA 1560  
Db  
1501 GTCTGGAATAGAAAGGCTACAGAAAGTGTGTTGGACAGTGGGTGTCGAGAGTACTCTA 1560  
QY  
1561 AACATGAGTGAAGCCATAAAGGTTCTCGTCAAGTCTGTCATACAAATTCCTTTGGC 1620  
Db  
1561 AACATGAGTGAAGCCATAAAGGTTCTCGTCAAGTCTGTCATACAAATTCCTTTGGC 1620  
QY  
1621 ACATCTGTGAGAGATCCCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCCAAGACAAC 1680  
Db  
1621 ACATCTGTGAGAGATCCCTTTAAACTCTCCAGGCCCTTCCCTTTTCATCCCAAGACAAC 1680  
QY  
1681 ATCTCATCTATGCAACAATGTTGTTGTCCTCTCTCATATGTCGTTTAAACCTGCTA 1740  
Db  
1681 ATCTCATCTATGCAACAATGTTGTTGTCCTCTCTCATATGTCGTTTAAACCTGCTA 1740  
QY  
1741 ATTTGTACAAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
Db  
1741 ATTTGTACAAAGTACAAAAGCAATTTAGGTATGAAAGCCAGCTACAGATGGTACAGGTG 1800  
QY  
1801 ACCGGCTCTCAGATATAGTACTTCTACGTTGATTTACAGAAATATGAATATGATCTC 1860  
Db  
1801 ACCGGCTCTCAGATATAGTACTTCTACGTTGATTTACAGAAATATGAATATGATCTC 1860  
QY  
1861 AATGGGAGTTTCCAGAGAAAATTTAGATTTGGGAAGTACTAGGATCAGGTGCTTTT 1920  
Db  
1861 AATGGGAGTTTCCAGAGAAAATTTAGATTTGGGAAGTACTAGGATCAGGTGCTTTT 1920  
QY  
1921 GGAAGAGTATGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGAGGTCTCAATCCAGGTT 1980  
Db  
1921 GGAAGAGTATGATGAACGCAACAGCTTATGGAATTTAGCAAAACAGAGGTCTCAATCCAGGTT 1980  
QY  
1981 GCGGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAAGAA 2040  
Db  
1981 GCGGTCAAAATGCTGAAAGAAAAGCAGACAGCTCTGAAAGAGAGGCACTCATGTCAAGAA 2040  
QY  
2041 CTCAGATGATGATGCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGCGTGC 2100  
Db  
2041 CTCAGATGATGATGCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGCGTGC 2100  
QY  
2101 ACATCTCAGGACCAATTTACTTGAATTTTGAATCTGTTGCTATGTTGATCTCTCAAC 2160  
Db  
2101 ACATCTCAGGACCAATTTACTTGAATTTTGAATCTGTTGCTATGTTGATCTCTCAAC 2160  
QY  
2161 TATCTAAGAGTAAAAGAGAAAATTTACAGGACTTTGGACAGAGATTTTTCAGGAACAC 2220  
Db  
2161 TATCTAAGAGTAAAAGAGAAAATTTACAGGACTTTGGACAGAGATTTTTCAGGAACAC 2220  
QY  
2221 AATTTTCAGTTTTCACCCACTTTTCCAAATCATCTCCAAATTCAGCATGCTGTTCAAGA 2280  
Db  
2221 AATTTTCAGTTTTCACCCACTTTTCCAAATCATCTCCAAATTCAGCATGCTGTTCAAGA 2280  
QY  
2281 GAAGTTACATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTAC 2340  
Db  
2281 GAAGTTACATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTAC 2340  
QY  
2341 TCTGAAGATCAAAATTCGAATATGAAACCAAAAGGCTGGAAGAGAGGAGGACTTCAAT 2400  
|||||

Db  
2341 TCTGAAGATGAAATTTGAATATGAAAACCAAAAAGGCTGGAAGAGAGGAGGACTTGAAT 2400  
QY  
2401 GTGCTTACATTTGAAGATCTTCTTGTTCGATATCAAGTTGCGCAAGGAATGGAATTT 2460  
Db  
2401 GTGCTTACATTTGAAGATCTTCTTGTTCGATATCAAGTTGCGCAAGGAATGGAATTT 2460  
QY  
2461 CTGGAATTTAAGTTCGTGTTTCACAGACACTGGCCGCCAGGAACGTCTCTGTCACCCAC 2520  
Db  
2461 CTGGAATTTAAGTTCGTGTTTCACAGACACTGGCCGCCAGGAACGTCTCTGTCACCCAC 2520  
QY  
2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTATTCACCAAC 2580  
Db  
2521 GGGAAAGTGGTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGAGTATTCACCAAC 2580  
QY  
2581 TATGTTGTGACGGGCAATGCGCGTCTGCTGTAAATGATGGCCCCCGAAGACCTGTTT 2640  
Db  
2581 TATGTTGTGACGGGCAATGCGCGTCTGCTGTAAATGATGGCCCCCGAAGACCTGTTT 2640  
QY  
2641 GAAGCATCTACACCATTAAGAGTATGCTGGTCAATATGGAATATTTACTGTGGAAATC 2700  
Db  
2641 GAAGCATCTACACCATTAAGAGTATGCTGGTCAATATGGAATATTTACTGTGGAAATC 2700  
QY  
2701 TTCTCACTTGTGTGAATCCTTACCCTGGCATTCGGTTGATGCTTAACCTTCTACAAACATG 2760  
Db  
2701 TTCTCACTTGTGTGAATCCTTACCCTGGCATTCGGTTGATGCTTAACCTTCTACAAACATG 2760  
QY  
2761 ATTCAAATGATTTAAATGATGACGCAATTTTATGTACAGAGAATAATATACATATA 2820  
Db  
2761 ATTCAAATGATTTAAATGATGACGCAATTTTATGTACAGAGAATAATATACATATA 2820  
QY  
2821 ATGCAATCTCTCGGCTTTTGAATCAAGGAACGGCCATCCTTCCCTAAATTTGACTTCG 2880  
Db  
2821 ATGCAATCTCTCGGCTTTTGAATCAAGGAACGGCCATCCTTCCCTAAATTTGACTTCG 2880  
QY  
2881 TTTTATGATGTACAGTGGCAGATGCAGAAAGCGATGTATCAAAATGTTGATGCGCT 2940  
Db  
2881 TTTTATGATGTACAGTGGCAGATGCAGAAAGCGATGTATCAAAATGTTGATGCGCT 2940  
QY  
2941 GTTTCGGAATGCTCTACACCTACAAAACAGGGACCTTTTACAGAGAGATGGAATTTG 3000  
Db  
2941 GTTTCGGAATGCTCTACACCTACAAAACAGGGACCTTTTACAGAGAGATGGAATTTG 3000  
QY  
3001 GGGTACTCTCTCGCAGGCTCAGGTCAAAGTTCGTAGAGGAACAAATTTAGTTTAAAG 3060  
Db  
3001 GGGTACTCTCTCGCAGGCTCAGGTCAAAGTTCGTAGAGGAACAAATTTAGTTTAAAG 3060  
QY  
3061 ACTTCATCTCCCTCACTATCCCTTAACAGGCTGTAGATTACCAAAACAAAGATTAATTTCA 3120  
Db  
3061 ACTTCATCTCCCTCACTATCCCTTAACAGGCTGTAGATTACCAAAACAAAGATTAATTTCA 3120  
QY  
3121 CACTAAAAGAAAATCTATTATCAACTGCTTCCACAGACTTTTCTCTAGAAGCGCTCT 3180  
Db  
3121 CACTAAAAGAAAATCTATTATCAACTGCTTCCACAGACTTTTCTCTAGAAGCGCTCT 3180  
QY  
3181 GCGTTTACTCTTGTTCAGAGGACTTTTGTAAATCAATCAATCTCTGTCACAGGCG 3240  
Db  
3181 GCGTTTACTCTTGTTCAGAGGACTTTTGTAAATCAATCAATCTCTGTCACAGGCG 3240  
QY  
3241 GAGGAGTGAATTAATGAATTTTATGGAGCATTTGATCTCATCCAGGCTTCTCTCAGGCG 3300  
Db  
3241 GAGGAGTGAATTAATGAATTTTATGGAGCATTTGATCTCATCCAGGCTTCTCTCAGGCG 3300  
QY  
3301 GCTTGAAGTGAATTTGACTACCTGAAAGTACAGTATATTTCTTGAATATACATAAAACAAAGC 3360  
Db  
3301 GCTTGAAGTGAATTTGACTACCTGAAAGTACAGTATATTTCTTGAATATACATAAAACAAAGC 3360  
QY  
3361 ATTTTGTAGGAGAGCTAATATGATTTTAAAGTCTATCTTTTAAATATATATGATAA 3420  
Db  
3361 ATTTTGTAGGAGAGCTAATATGATTTTAAAGTCTATCTTTTAAATATATATGATAA 3420  
QY  
3421 TTTTTCAGCTATTTAGTCAATATTTTATGGTGGGAATAAATTTCTACTACAGAAAAA 3480  
Db  
3421 TTTTTCAGCTATTTAGTCAATATTTTATGGTGGGAATAAATTTCTACTACAGAAAAA 3480



```
QY 1318 GAAATGATGATCCCAATTTACCAAAATGTTCCAGCTGAATATAGAAGGAAACCTCAA 1377
DB 1261 GAAATGATGATCCCAATTTACCAAAATGTTCCAGCTGATATAGAAGGAAACCTCAA 1320
QY 1378 GTCTCGCAGAACATCGGCAAGTCAGGCGTCTGTTTCGCGATGATACCATACCA 1437
DB 1321 GTCTCGCAGAACATCGGCAAGTCAGGCGTCTGTTTCGCGATGATACCATACCA 1380
QY 1438 TCTTGACCTGGGAAGTGTTCAGACAACTCTCCCAACTGCACAGAGAGATCACAGAA 1497
DB 1381 TCTTGACCTGGGAAGTGTTCAGACAACTCTCCCAACTGCACAGAGAGATCACAGAA 1440
QY 1498 GGAGTCTGGAATAGAAGGCTAACAGAAAAGTGTTCGAGAGGGTGTGAGCAGTACT 1557
DB 1441 GGAGTCTGGAATAGAAGGCTAACAGAAAAGTGTTCGAGAGGGTGTGAGCAGTACT 1500
QY 1558 CTAACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGCGATACAATTCCTTT 1617
DB 1501 CTAACATGAGTGAAGCCATAAAGGGTTCCTGGTCAAGTGTGCGATACAATTCCTTT 1560
QY 1618 GGCACATCTGTGAGACGATCCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAGAC 1677
DB 1561 GGCACATCTGTGAGACGATCCTTTTAACTCTCCAGGCCCTTCCCTTTCATCCAGAC 1620
QY 1678 AACATCTCATTTATGCAACAAATGTTGTTGTTCTCTCTTCAATGTCGTTTAAACCTG 1737
DB 1621 AACATCTCATTTATGCAACAAATGTTGTTGTTCTCTCTTCAATGTCGTTTAAACCTG 1680
QY 1738 CTAATTTGCAAGTACAAAAGCAATTTAGTATGAAGCAGCTACAGATGGTACAG 1797
DB 1681 CTAATTTGCAAGTACAAAAGCAATTTAGTATGAAGCAGCTACAGATGGTACAG 1740
QY 1798 GTGACCGCTCTCAGATATGATGATCTCTAGTGTGATTTGAGAAATGATGAT 1857
DB 1741 GTGACCGCTCTCAGATATGATGATCTCTAGTGTGATTTGAGAAATGATGAT 1800
QY 1858 CTCAAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAAGTACTAGATCAGGTGCT 1917
DB 1801 CTCAAATGGGAGTTTCCAAAGAGAAATTTAGAGTTTGGGAAGTACTAGATCAGGTGCT 1860
QY 1918 TTTGGAAGTATGATGAACCCACAGCTTATGAAATGAGAAACAGAGATCTCAATCCAG 1977
DB 1861 TTTGGAAGTATGATGAACCCACAGCTTATGAAATGAGAAACAGAGATCTCAATCCAG 1920
QY 1978 GTTGGCGTCAAAATGCTGAAAGAAAGCAGACAGCTCTGAAAGAGAGCAGCTCATGTCA 2037
DB 1921 GTTGGCGTCAAAATGCTGAAAGAAAGCAGACAGCTCTGAAAGAGAGCAGCTCATGTCA 1980
QY 2038 GAACCTCAAGATGATGACCCAGCTGGGAAGCCACGAGAATATTGTGAACCTGCTGGGGCG 2097
DB 1981 GAACCTCAAGATGATGACCCAGCTGGGAAGCCACGAGAATATTGTGAACCTGCTGGGGCG 2040
QY 2098 TGCACACTGTGAGGACCAATTTACTTGATTTTGAATGATGTTGCTGATGATCTTC 2157
DB 2041 TGCACACTGTGAGGACCAATTTACTTGATTTTGAATGATGTTGCTGATGATCTTC 2100
QY 2158 AACTATCTAAGAGTAAAGAGAAATTTACAGAGCTTGACAGAGATTTTCAGGAA 2217
DB 2101 AACTATCTAAGAGTAAAGAGAAATTTACAGAGCTTGACAGAGATTTTCAGGAA 2160
QY 2218 CACAAATTCAGTTTATACCCACTTTCCAAATCACATCCAAATTCAGCATGCTGGTTCA 2277
DB 2161 CACAAATTCAGTTTATACCCACTTTCCAAATCACATCCAAATTCAGCATGCTGGTTCA 2220
QY 2278 AGAGAGTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTT 2337
DB 2221 AGAGAGTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTT 2280
QY 2338 CACTCTGAAGTGAATTTGAATATGAARACCAAAAGGCTGGAAGAGAGAGGACTTG 2397
DB 2281 CACTCTGAAGTGAATTTGAATATGAARACCAAAAGGCTGGAAGAGAGAGGACTTG 2340
QY 2398 AATGTGCTTACATTTGAAGATCTTCTTTCGTTTGCATATCAAGTTGCCAAAGGATGGAA 2457
```

```
DB 2341 AATGTGCTTACATTTGAAGATCTTCTTGTGTTTGCATATCAAGTTGCCAAGGATGAA 2400
QY 2458 TTTCTGGAATTTAAGTCTGTTTACAGAGACCTGGCCGCGAGAACGTCCTTGTTCACC 2517
DB 2401 TTTCTGGAATTTAAGTCTGTTTACAGAGACCTGGCCGCGAGAACGTCCTTGTTCACC 2460
QY 2518 CACGGGAAGTGGTGAAGATATGAGTCTTGGATTGGCTCCAGATATCATCAGATGATCC 2577
DB 2461 CACGGGAAGTGGTGAAGATATGAGTCTTGGATTGGCTCCAGATATCATCAGATGATCC 2520
QY 2578 AACTATCTTGTTCAGGGCAATGCCGCTGCTGCTGTAATAATGGATGGCCCCGGAAGCCCTG 2637
DB 2521 AACTATCTTGTTCAGGGCAATGCCGCTGCTGCTGTAATAATGGATGGCCCCGGAAGCCCTG 2580
QY 2638 TTTGAAGCATCTACACCATTAAGAGTGTCTGGTCATATGAATATTTACTGTGGAA 2697
DB 2581 TTTGAAGCATCTACACCATTAAGAGTGTCTGGTCATATGAATATTTACTGTGGAA 2640
QY 2698 ATCTTCTCACTTGTGTGAATCCTTACCCTGGCATTCGGTTCGATGCTAACTTCTACAA 2757
DB 2641 ATCTTCTCACTTGTGTGAATCCTTACCCTGGCATTCGGTTCGATGCTAACTTCTACAA 2700
QY 2758 CTGATTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATT 2817
DB 2701 CTGATTCAAAATGGATTTAAAATGGATCAGCCATTTTATGCTACAGAAAGAAATATACATT 2760
QY 2818 ATAATGCAATCTGCTGGGCTTTTGTACTCAAGAAAGCGCCATCTTCCCTTAATTTGACT 2877
DB 2761 ATAATGCAATCTGCTGGGCTTTTGTACTCAAGAAAGCGCCATCTTCCCTTAATTTGACT 2820
QY 2878 TCCTTTTAGATGTGCTGAGCTGGCAGATGCAGAGAACGCGATGATCAGATGTGGATGGC 2937
DB 2821 TCCTTTTAGATGTGCTGAGCTGGCAGATGCAGAGAACGCGATGATCAGATGTGGATGGC 2880
QY 2938 CGTGTTCGGAATGCTCCTCACACTACACCTACAAAACAGGCGACCTTTTCAGCAGAGATGGAT 2997
DB 2881 CGTGTTCGGAATGCTCCTCACACTACACCTACAAAACAGGCGACCTTTTCAGCAGAGATGGAT 2940
QY 2998 TTGGGCTACTCTCTCCGAGGCTCAGTCTGAGATTCGAGAGAACAAATTTAGTTTAA 3057
DB 2941 TTGGGCTACTCTCTCCGAGGCTCAGTCTGAGATTCGAGAGAACAAATTTAGTTTAA 3000
QY 3058 AGCACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTTACCAAAACAGATTAATTT 3117
DB 3001 AGCACTTCATCCCTCCACCTATCCCTAACAGGCTGTAGATTTACCAAAACAGGTTAATTT 3060
QY 3118 CATCACTAAAAGAAATCTATTATCAACTGCTGCTTCACAGACTTTTCTCTAGAGCCG 3177
DB 3061 CATCACTAAAAGAAATCTATTATCAACTGCTGCTTCACAGACTTTTCTCTAGAGAGCG 3120
```

## RESULT 4

US-09-919-408-1

; Sequence 1, Application us/09919408

; Patent No. US20020072077A1

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor R.

; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL

; RECEPTORS AND THEIR LIGANDS

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Imclone Systems Incorporated

; STREET: 180 Varick Street

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10014

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25



CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/919,408  
FILING DATE: 31-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/977,451  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3453 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 112...3006  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 31...111  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 31...3009  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Query Match. 63.6%; Score 2227.2; DB 10; Length 3453;  
Best Local Similarity 79.6%; Pred. No. 0;  
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;  
QY 35 GGGAGCCCGGGTCCGGAGGCATCGCGGCTTGGC---GCCGAGCGGGGACCGTGC 91  
DB 8 GGCTACCGCGCGTCCGGAGGCATCGCGGCTTGGCGAGCGAGCGCGCGGCTGC 67  
QY 92 CGCTGCTGCTGTTTTCGCAATGATATTGGGACTATTACAAATCAAGATCTGCCTG 151  
DB 68 TGTGCTGTTGTTTGTCTAGTATGATCTTGAGACCGTTTACAAACAGACCGCTGCTG 127  
QY 152 TGATCAAGTGTGTTTAAATCAATCAATGAACAAATGATTCATAGTGGGGAAGTCATCAT 211  
DB 128 TGATCAAGTGTGTTTAAATCAATCAATGAACAAATGATTCATAGTGGGGAAGTCATCAT 187  
QY 212 CATATCCCATGTTATCAAGATCCCGGAGACCTCGGGGTGCGCTTGAGACCCCGAGAGCT 271  
DB 188 CGTACCAGATGGTCCGAGGATCCCGAGAGACCTCCAGTGTACCCCGAGGCGCGAGAGTG 247  
QY 272 CAGGAGAGTGATACGAAGCTCCCGCTGTGGAGTGGATGTATCTGCTTCCATCACACTGC 331

DB 248 AAGGACGGTATATGAAGCGCCACCGTGGAGGTGGCGAGTCTGGGTCCATCACCCTGC 307  
QY 332 AAGTCTGGTCGATGCCCGAGGGAACATTTCTCTCTGGGTCTTTAAGCACAGCTCCC 391  
DB 308 AAGTGCAGCTCGCCACCGCCAGGGACCTTCTCTCTGGGTCTTTAAGCACAGCTCCC 367  
QY 392 TGAATTGCCAGCCACATTTTGTATTTACAAACAGAGGAGTGTCTTCCATGGTCATTTGA 451  
DB 368 TGGGCTGCCAGCCGACCTTTGATTTTACAAACAGAGGAATCGTTTCCATGGCCATCTGA 427  
QY 452 AAATGACAGAAACCAAGCTGGAGAAATACCTACTTTTATTTCAGAGTGAAGTACCAAT 511  
DB 428 ACGTGACAGAGCCCGAGGAGGAGTACCTACTCCATATTCAGAGCGAAGCGCAACT 487  
QY 512 ACACAATATGTTTACAGTGTAGTATAAGAAATACCTGCTTTTACACATTAAGAAGACCTT 571  
DB 488 ACACAGTACTGTTACAGTGAATGTAAGAGATACACAGCTGTATGTCTAAGAGACCTT 547  
QY 572 ACTTTAGAAAATGGAACACAGGAGCGCCCTGGTCTGTCATATCTGAGACGTTTCCAGAGC 631  
DB 548 ACTTTAGGAAGATGAAAACACAGGATGCACTGCTCTGCATCTCCGAGGCTGTTCCGAGC 607  
QY 632 CGATCGTGAATGGGTGCTTTTCGATTACAGGGGGGAAAGCTGTAAGAGAAAGTCCAG 691  
DB 608 CCACTGTGGAGTGGGTGCTCTGCAGCTCCACAGGGAAGCTGTAAGAGAAAGGCGCTG 667  
QY 692 CTGTTGTTAAAAGGAGGAAAAGTCTTCATGAATTTTGGGAGCGGACATGAAGTGCT 751  
DB 668 CTGTTGTCAGAAAGGAGGAAAAGTCTTCATGAGTTGTTCCGAGACAGATCAGATGCT 727  
QY 752 GTGCCAGAAATGAATGGGAGGGAATGCACAGGCTGTTTCAACAATAGATCTAAATCAA 811  
DB 728 GTGCTAGAAATGCATCGGCGCGAATGCACCAAGCTGTTCCACATAGATCTAAACAGG 787  
QY 812 CTCCTCAGACCATTTGCCACATTTATTTCTTAAAGTAGGGGAACCTTTATGATAAGT 871  
DB 788 CTCCTCAGACCATTTGCCACATTTATTTCTGAAAGTGGGGGAACCTTTGATCAGT 847  
QY 872 GCAAGCTGTTTCATGTGAACCATGGAATTCGGGCTCACCTGGGAATTAGAAAACAAAGC 931  
DB 848 GTAAGCCCATCCATGTGAACCATGGAATTCGGGCTCACCTGGAGCTGGAAGACAAAGC 907  
QY 932 TCAGAGGAGGCAACTACTTTGAGATGAGTACCTATTCAACAACAGACACTATGATACGGA 991  
DB 908 TGCAGGAGGCGAGTACTTTGAGATGAGTACCTACTCCACAACAGACCATGATTCGGA 967  
QY 992 TTCTGTTGCTTTTGTATCATCAGTGCAGAAAGACGACCGGATCTACACTTGTTCCT 1051  
DB 968 TTCTCTTGGCCTTTGTGCTTCCGTGGGAGGAGACGACCGGATATTACACTGCTCTT 1027  
QY 1052 CTTCAAGCATCCAGTCAATCAGCTTTGGTTACCATCTAGGAAAGGATTTATAATG 1111  
DB 1028 CTTCAAGCATCCCGAGGAGTACGCTGAGTGGTGACCATCTCAGAAAAGGGTTTATAACG 1087  
QY 1112 CTACCAATTCAGTGAAGATTTGAAATTCACCAATATGAAGAGTGTGTTTTCTGTCA 1171  
DB 1088 CTACCACTCGCAAGAGAGTATGAAATTCACCCGTACGAAAAGTCTGCTTCTCAGTCA 1147  
QY 1172 GGTGTTAAAGCTACCCACAAATCAGATGATGAGTGGACCTTCTCTCGAAAATCATTCCT 1231  
DB 1148 GGTGTTAAAGGCTACCCACGAAATCCGATCGCATGGATCTTCTCTCAGGCTCATTTCT 1207  
QY 1232 GTGAGCAAAAGGGTCTTTGATAACGATACAGCATATCCAACTTTTGAATCATAGCAACC 1291  
DB 1208 GTGAACAGAGAGGCTCGAGGATGGGTACACATATCTAAATTTTGGATCATAGAACA 1267  
QY 1292 AGCCAGGAGATATATATTTCCATCGAATAATGATGATGCCAATTTTACCATAATGTTCA 1351  
DB 1268 AGCCAGGAGATACATATTTCTATCGAATAATGATGACGCCGCTTACCAAAAATGTTCA 1327  
QY 1352 CGCTGAATATTAAGAGGAAACCTCAAGTGTGCGAGAGCATCGGCAAGTACAGGCTCT 1411

Db 1328 CGCTGAATATAAGAAAGAAACCTCAAGTGTCTAGCAATGCTCAGCCAGCCAGCGCTCCT 1387  
QY 1412 GTTTCCTCGGATGGATACCCATTACCATTCTGGACCTGGGAAGAAGTGTTCAGACAAGTCTC 1471  
Db 1388 GTTCTCTGATGCTACCCGCTACCCTCTTGGACCTGGGAAGAAGTGTTCGGGCAAAATCTC 1447  
QY 1472 CCAACTGCACAGAGAAGATCACAGAAGGAGTCTGGAATAGAAAGGCTAACAGAAAAGTGT 1531  
Db 1448 CCAATTGCACGGAGAAATCCCAAGAGGAGTTTGGAAATAAAAAGGCTAACAGAAAAGTGT 1507  
QY 1532 TTGGACAGTGGGTCTGCAGCAGTACTCTAAACATGAGTGAAGCCATAAAGGGTTCCTGG 1591  
Db 1508 TTGGCCAGTGGGTCTGCAGCAGTACTCTAAATATGAGTGGGCGGGAAGAGGCTTCCTGG 1567  
QY 1592 TCAAGTGTCTGCATACAAATTCCTTGGCACATCTTGTGAGACGATCCCTTTTAAACTCTC 1651  
Db 1568 TCAATGTCTGCGTACAAATCTATCGGCACGCTCTTGGCAACCATCTTTTAAACTCAC 1627  
QY 1652 CAGGCCCTTCCTTTCATCCAAAGCAACATCTCATTTCTATGCAACAATTTGTTTGTG 1711  
Db 1628 CAGGCCCTTCCTTTCATCCAAAGCAACATCTCTCTATGCGACCATTTGGGCTCTGTC 1687  
QY 1712 TCCTCTTCAATGTCTTTTAAACCTGCTAAATTTGTCTACAAGTACAAAAAGCAATTTAGGT 1771  
Db 1688 TCCTCTTCAATGTCTTTTAAACCTGCTAAATTTGTCTACAAGTACAAAAAGCAATTTAGGT 1747  
QY 1772 ATGAAAGCCAGCTACAGATGGTACAGTGTGACCGGCTCCTCAGATATGAGTACTTCTACG 1831  
Db 1748 ACGAGAGTCAGCTGCAGATGATCCAGTGACTGGCCCCCTGGATACGAGTACTTCTACG 1807  
QY 1832 TTGATTTTCAGAAATATGAAATATGATCTCAAAATGGGAGTTTCCAAAGAAAATTTAGAGT 1891  
Db 1808 TTGACTTTCAGGACTATGAATATGACTTAAAGTGGGAGTTTCCCGAGAGAACTTAGAGT 1867  
QY 1892 TTGGGAAGTACTAGGATCAGTGTCTTTTGGAAAAGTGTGACGCAACAGCTTATGGAA 1951  
Db 1868 TTGGGAAGTCTGGGGTCTGGGCTTTCGGGAGGAGTGTGAACGCCACGGCTATGGGA 1927  
QY 1952 TTAGCAAAAACAGGAGTCTCAATCCAGTTTCCGCTCAAAATGCTGAAAGAAAACAGACAGA 2011  
Db 1928 TTAGTAAACGGAGTCTCAATTCAGTGGGCTGAAGTGTAAAGAGAAAGCTTGACA 1987  
QY 2012 GCTCTGAAAGAGAGGACACTCATGTCCAGAACTCAAGATGATGACCCAGCTGGGAAGCCAG 2071  
Db 1988 GCTGTGAAAAGAGGCTCATGTCCGAGCTCAAAATGATGACCCACCTGGGACACCATG 2047  
QY 2072 AGAATATGTGAACCTGTGGGGGCTGCACACTGTCAGACCAATTTACTTGATTTTGG 2131  
Db 2048 ACAACATCGTGAATCTGTGGGGGCTGCACACTGTCAGGGCCAGTGTACTTGATTTTGG 2107  
QY 2132 AATACTGTGTGCTATGCTGATCTTCTCAACTATCTAAGAAAGTAAAGAGAAAATTTTACA 2191  
Db 2108 AATATTTGTGCTATGCTGACCTCTCACTACTTAAGAAAGTAAAGAGAAAGTTTACA 2167  
QY 2192 GGACTTGGACAGAGATTTTCAAGGAACACAAATTTACGTTTAAACCCACTTTCCAATCAC 2251  
Db 2168 GGACATGGACAGAGATTTTAAAGAACATAATTTACGTTCTTACCTACTTTCAGGGCAG 2227  
QY 2252 ATCCAAATTCAGCATGCTGCTTCAAGAGAGTTCAGATACACCCGGGACTCGGATCAAA 2311  
Db 2228 ATTCAAATTCAGCATGCTGCTTCAAGAGAGTTCAGTATACACCCGGCTTGGATCAGC 2287  
QY 2312 TCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAATTTGAATGAACCAAA 2371  
Db 2288 TCTCAGGGTTCATGGGAATTCATTTCACTCTGAAGATGAATTTGAATGAACCAAA 2347  
QY 2372 AAAGGCTG-----GAAAGAGAGGAGGACTTGAATGTGCTTACATTTGAAGATCTCTTT 2425  
Db 2348 AGAGCTGGCAAGAGAGAGGAGAGATTTGAACCTGCTGACGTTTGAAGACCTCCTTT 2407  
QY 2426 GCTTTGCAATATCAAGTGTGCCAAAGGAATTTCTGGAATTTAAGTCTGTGTTTACA 2485  
Db 2408 GCTTTGCGTACCAAGTGGCCAAAGGATTCCTCGAGTTCAAGTCTGTGTGCCACA 2467

QY 2486 GAGACCTGGCGCCAGGAAGCTGCTTGTCTACCCACCGGAAGTGTGAAGATATGTGACT 2545  
Db 2468 GAGACCTGGCGCCAGGAAGTGTGTGTGTACCCACCGGAAGTGTGAAGATATGTGACT 2527  
QY 2546 TTGGATTGGCTCGAGATATCATGAGTGAATTCCTCAACTATGTTGTGAGGGCAATGCCGCTC 2605  
Db 2528 TTGGACTGGCGCCAGACATCTTGTGAGCGACTCCAGCTACGTCGTGAGGGCAACGACGGC 2587  
QY 2606 TGCTGTAAATGGATGGCGCCCGAAAGCCCTGTTTGAAGGCATCTACACCAATTAAGAGTG 2665  
Db 2588 TGCCGGTGAAGTGGATGGCACCAGAGACTTATTTGAAGGGATCTACACAATCAAGAGTG 2647  
QY 2666 ATGCTGTGTCATATGAATATTAAGTGTGGGAAATCTTCTCAGTGTGGTGAATCCTTACC 2725  
Db 2648 ACCTGTGGTCTTACGGCATCTCTCTGGGAGATATTTTACCTGGGTGAACCCCTTACC 2707  
QY 2726 CTGGCATTTCCGGTTGATGCTTAACCTCTTACAACTGATTCAAAATGGATTTTAAATGGATC 2785  
Db 2708 CTGGCATTTCTGTGCGAGCTAACTTCTTATAAAGTATTCAGAGTGGATTTTAAATGGAGC 2767  
QY 2786 AGCATTTTATGCTACAGAGAATAATATACATTTAATGCAATCTCTGCTGGGCTTTTGACT 2845  
Db 2768 AGCATTTCTATGTCACAGAGGATATTAATTTGTAATGCAATCTCTGCTGGGCTTTTGACT 2827  
QY 2846 CAAGGAAACGGCCATCTTCCCTAAATTTGACTTCTGTTTGTAGGATGTGAGTGGCAGATG 2905  
Db 2828 CAAGGAAACGGCCATCTTCCCTAACTGACTTCTTATTTTAGGATGTGAGTGGCAGAGG 2887  
QY 2906 CAGAAGAAGCGATGTATCAGATGTGGATGGCGCTGTTTGGGATGTCTCCACACCTTACC 2965  
Db 2888 CAGAAGAAG-----CATGTATCAGAACATCCATCTTACC 2924  
QY 2966 ABAACAGGCGACCTTTCAGCAGAGAGATGATTTGGGGCTACTCTCTCCGAGGCTCAGG 3025  
Db 2925 ABAACAGGCGCGCTTTCAGCAGAGAGGCGGCTCAGAGC-CCAGTCCGACACGCCAGG 2983  
QY 3026 TCCAAGATTCGTAGAGAACAAATTTAGTCTTAAAGGACTTATCCCTCCACCTATCCCTAA 3085  
Db 2984 T-GAAGATTCACAGAGAAGAAAGTTAGCGAGGAGGCGCTTGGACCCCGC----CACCTAG 3038  
QY 3086 CAGGCTGTAGATTTACCAAAAACAAGATTAATTTTATCATCTATAAAGAAAATCTATTATCAAC 3145  
Db 3039 CAGGCTGTAGACCGCAGAGCCAAAGATTAGCTTCGCTCT--GAGGAAGCGCCCTACAGCG 3096  
QY 3146 TGCTGCTTCACAGACTTTTCTCTAGAACCGCTGCGGTTTACTCTTGTGTTTTCAGAGGA 3205  
Db 3097 CGTTGCTTCGCTGAGCTTTTCTCTAGATGCTGTGCTGCAATTACTC-----CAAAGTGA 3149  
QY 3206 CTTTTCFAAATCAAAATCATCTCTCACAGGCGAGGAGGAGCTCATTAATGAACCTTTATTG 3265  
Db 3150 CTTCTATAAATCAAACTCTCTCTCGCACAGCGGGAGAGCCCAATAATAGACTTGTGTGG 3209  
QY 3266 GAGCATTTGATCTGCATCCCAAGCGCTTCTCAGCGCGGCTTTCAGTGAATTTGTACTCAAG 3325  
Db 3210 TGAGCCCGCTTACCTGGGGCTTTCCAGCAGCTTCAGGGGAAAGCCATGTATCTGAAA 3269  
QY 3326 TACAGTATATCTTGTAAATACATAAAACAAA-----AGCATTTTGTCAAGGAGAGC 3378  
Db 3270 TATAGTATATCTTGTAAATACCTGTAACAAACAAACCCGTTTTTTGCTAAGGAAAGC 3329  
QY 3379 TAATATGATTTTTT--AAGTCTATGTTTAAATAATATGTAATTTTTCAGCTATTATG 3436  
Db 3330 TAAATATGATTTTTTAAATAATCTATGTTTAAATACTATGTAACCTTTTTCATCTATTAG 3389  
QY 3437 TGATATATTTTATGGTGGGAATAAATTTCTACTACAGAAAAAATAAATAAATAAATAA 3496  
Db 3390 TGATATATTTTATGGATGGAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3449  
QY 3497 AAAA 3500  
Db 3450 AAAA 3453

RESULT 5  
US-09-872-136-1  
Sequence 1, Application US/09872136  
Patent No. US20020119545A1  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Imclone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/872,136  
FILING DATE: 01-Jun-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/208,786  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/09/021,324  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/07/977,451  
FILING DATE: 1992-11-19  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3453 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 112...3006  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 31...111  
FEATURE:

NAME/KEY: CDS  
LOCATION: 31...3009  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-872-136-1  
Query Match  
Best Local Similarity 63.6%; Score 2227.2; DB 10; Length 3453;  
Matches 2775; Conservative 0; Mismatches 653; Indels 56; Gaps 10;  
35 GGGACCCCGGCTCCGGAGGCCATCGCGCGTGGC---GGCGACGCGGCGCACCGCTGC 91  
8 GGCTACCGCGGCTCCGGAGGCCATCGCGCGTGGCAGCGAGCCGCGCGGCTGC 67  
92 CGCTCGTGTGTTTCTTCAATGATATTTGGGACTATTACAAATCAAGATCTGCCCTG 151  
68 TGCTGCTGTTGTTTGTCAAGTAATGTTTGTAGACCGTTACAAACCAAGACCTGCCCTG 127  
152 TGATCAAGTGTGTTTAAATCAATATAGAAACAATGATTCATCAGTGGGGAAGTCATCAT 211  
128 TGATCAAGTGTGTTTAAATCAATGATGAGAAATGGCTCATCAGCGGGAAGCCATCAT 187  
212 CATATCCCATGGTATCAGAAATCCCGGAAGACCTCGGGTGTGGTGTGAGACCCAGACT 271  
188 CGTACCGAATGTCGAGGATCCCGAAGAGACTCCAGTGTACCCGAGCGCGCAGAGTG 247  
272 CAGGACAGTGTACGAAGCTGCCGCTGTGGAAGTGGATCTATCTGCTCCATCACACTGC 331  
248 AAGGACGCTATATGAAGCGGCCACCGTGGAGGTGGCCGAGTCTGGGTCCATCACCCTGC 307  
332 AAGTGTGTCGATGCCCGGAGGAACTTCTGCTCTCTGGGTCTTTAAGCACAGCTCCC 391  
308 AAGTGCAGTCTGCCCGGAGGACCTTCTGCTCTGCTCTGGGTCTTTAAGCACAGCTCCC 367  
392 TGAATTGCCAGCACATTTTGATTTACAAACAGAGAGTGTGTTTCCATGGTCAATTTGA 451  
368 TGGGCTGCCCGGACACTTTGATTTACAAACAGAGAAATCGTTTCCATGCCATCTTGA 427  
452 AAATGACAAACCAAGCTGGAGATACCTACTTTTATTCAGAGTGAAGCTACCAATT 511  
428 ACCTGACAGAGACCCAGGAGGAGATACCTACTTCCATATTCAGAGGAAAGCGGCCAACT 487  
512 ACACATATTTGTACAGTGTATAGAAATACCTGCTTTACACATTAAGAAGACCTT 571  
488 ACACAGTACTGTTACAGTGAATGTAGAGATACACAGCTGTATGCTAAGGAGACCTT 547  
572 ACTTTAGAAAATGAAAACAGGACGCCCTGCTGTCATATCTGAGAGGCTTCCAGAGC 631  
548 ACTTTAGGAAGATGAAAACAGGATGCACTGCTCTGCTCTCGAGGGTGTTCGGAGC 607  
632 CGATCGTGAATGGTGTCTTGGATTCACAGGGGGAAGCTGTAAAGAAAGAGTCCAG 691  
608 CCACTGTGGAGTGGTGTCTGCGAGCTCCACAGGAAAGCTGTAAAGAAAGAGGCGCTG 667  
692 CTGTTGTTAAAAGAGGAGAAAAGTGTCTCATGAATTTATTTGGGACGACATAGGTGCT 751  
668 CTGTTGTGAGAAAGGAGAAAAGTGTCTCATGATTTGTTGCGNACAGACATCAGATGCT 727  
752 GTCCAGAAATGAATGGGAGGGAATGCACAGGCTGTTTCACAATAGATCTAATCAAA 811  
728 GTCTAGAAATGCACTGGGCGCGAATGCACCAAGCTGTTCAACCATAGATCTAAACAGG 787  
812 CTCTCAGACCATTTGCCACAAATTTCTTAAAGTAGGGGACCCCTTATGGATAGGT 871  
788 CTCTCAGAGCACACTGCCCCAGTTTATCTGAAAGTGGGGGAACCCCTTGTGGATCAGT 847  
872 GCAAGCTGTTCATGTGAACCATGGATTCGGGCTCACCTGGGAATTAGAAAACAAAGCAC 931  
848 GTAAGGCCATCCATGTGAACCATGGATTCGGGCTCACCTGGGAGCTGGAGACAAAGCCC 907  
932 TCGAGGAGGGCACTACTTTGAGATGAGTACCTATTCAACAAACAGAACTATGATACGGA 991  
908 TGGAGGAGGGCAGCTACTTTGAGATGAGTACCTACTCCACAAACAGGACCATGATTCGGA 967



Db 3097 CGTTGCTCGGAGCTTTCTCTAGATGCTGCTGCCATTACTC-----CAAGTGA 3149  
QY 3206 CTTTGTAAATCAATCATCTCTCAAGACGAGGAGCTGATGAATGAATTTATTG 3265  
Db 3150 CTTCTATAAATCAAACTCTCTCGCACAGCGGGAGAGCCAAATAATGAGACTTCTTGG 3209  
QY 3266 GAGCAATGATCGCATCAAGCCCTTCTCAGCGCGCTTGAGTGAATTTGTGTACCTGAAG 3325  
Db 3210 TGAGCCGCGCTACCTCGGGGCTTTCCACGAGCTTGAGGGGAAGCCATGATCTGAA 3269  
QY 3326 TACAGTATATCTTGTAAATACATAAACA-----AGCATTTTGTAGGAGAAGC 3378  
Db 3270 TATAGTATATCTTGTAAATACGTAACAAACCAACCCGTTTGTAGGGAAGC 3329  
QY 3379 TAATATGATTTT---AAGCTATGTTTAAATAATATGTAATTTTTCAGCTATTAG 3436  
Db 3330 TAAATATGATTTTAAATAATCTATGTTTAAATACTATGTAATTTTTCATCTATTAG 3389  
QY 3437 TGATATATTTATGGTGGGAATAAATTTCTACTACAGAAAAA-----AAAA 3496  
Db 3390 TGATATATTTATGATGGAATAAATTTCTACTGTAATAA-----AAAA 3449  
QY 3497 AAAA 3500  
Db 3450 AAAA 3453

RESULT 6  
US-09-796-692-7538  
; Sequence 7538, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09/796,692  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; NUMBER OF SEQ ID NOS: 9597  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7538  
; LENGTH: 645  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-796-692-7538

Query Match 18.1%; Score 633; DB 9; Length 645;

Best Local Similarity 99.8%; Pred. No. 1.3e-152;  
Matches 644; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 2059 CTGGGAAGCCACGAGAAATATTGTGAACCTGCTGGGGCGTGCACACTGTACAGACCAATT 2118  
Db 1 CTGGGAAGCCACGAGAAATATTGTGAACCTGCTGGGGCGTGCACACTGTACAGACCAATT 60  
QY 2119 TACTTGATTTTGAATACTGTTGCTATGTTGCTATGTTGCTCAACTATCTAAGAAGTAAAGA 2178  
Db 61 TACTTGATTTTGAATACTGTTGCTATGTTGCTATGTTGCTCAACTATCTAAGAAGTAAAGA 120  
QY 2179 GAAAAATTTTACAGAGACTTGGACAGAGATTTCAGGAACACAAATTTTACGTTTTTACCCC 2238  
Db 121 GAAAAATTTTACAGAGACTTGGACAGAGATTTCAGGAACACAAATTTTACGTTTTTACCCC 180  
QY 2239 ACTTTCCAATCACATCCAAATTCACAGATGCTGCTGTTTCAAGAGAAGTTCAGATACACCCG 2298  
Db 181 ACTTTCCAATCACATCCAAATTCACAGATGCTGCTGTTTCAAGAGAAGTTCAGATACACCCG 240  
QY 2299 GACTCGATCAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAATGAA 2358  
Db 241 GACTCGATCAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGATGAATGAA 300  
QY 2359 TATGAAACCAAAAAAGGCTGGAAGAGAGAGACTTGAATGTGCTTACATTTGAAGAT 2418  
Db 301 TATGAAACCAAAAAAGGCTGGAAGAGAGAGACTTGAATGTGCTTACATTTGAAGAT 360  
QY 2419 CTTCTTTGCTTTGCATATCAAGTTGCCAAAGAAATGGAATTTCTGGAATTTAAGTCGTGT 2478  
Db 361 CTTCTTTGCTTTGCATATCAAGTTGCCAAAGAAATGGAATTTCTGGAATTTAAGTCGTGT 420  
QY 2479 GTTCACAGAGACCTGGCGCCAGAGAGCTGCTGTACCCACGGGAAAGTGGTGAAGATA 2538  
Db 421 GTTCACAGAGACCTGGCGCCAGAGAGCTGCTGTACCCACGGGAAAGTGGTGAAGATA 480  
QY 2539 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTCACACTATGTTGTGAGGGCAAT 2598  
Db 481 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGATTCACACTATGTTGTGAGGGCAAT 540  
QY 2599 GCCCGTCTGCTCTAAATGATGGCCCGCCAGAGAGCTGTTGAGAGGCTG-TTTGAAGGCATCTACACCAT 2657  
Db 541 GCCCGTCTGCTCTAAATGATGGCCCGCCAGAGAGCTGTTGAGAGGCTGTTGAGAGGCATCTACACCAT 600  
QY 2658 TAAGAGTGATGCTGCTGATATGATATGATATGATGATGATGATGATGATGATGATGAT 2702  
Db 601 TAAGAGTGATGCTGCTGATATGATATGATATGATGATGATGATGATGATGATGATGAT 645

RESULT 7  
US-10-040-862-7538  
; Sequence 7538, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther  
; FILE REFERENCE: 014058-0135200S  
; CURRENT APPLICATION NUMBER: US/10/040,862  
; CURRENT FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28

;; PRIOR APPLICATION NUMBER: US 60/200,999  
;; PRIOR FILING DATE: 2000-05-01  
;; PRIOR APPLICATION NUMBER: US 60/202,084  
;; PRIOR FILING DATE: 2000-05-04  
;; PRIOR APPLICATION NUMBER: US 60/206,201  
;; PRIOR FILING DATE: 2000-05-22  
;; PRIOR APPLICATION NUMBER: US 60/218,950  
;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: US 60/222,903  
;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: US 60/223,416  
;; PRIOR FILING DATE: 2000-08-04  
;; PRIOR APPLICATION NUMBER: US 60/223,378  
;; PRIOR FILING DATE: 2000-08-07  
;; PRIOR APPLICATION NUMBER: US 09/796,692  
;; PRIOR FILING DATE: 2001-03-01  
;; NUMBER OF SEQ ID NOS: 10467  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 7538  
;; LENGTH: 645  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-040-862-7538

Query Match 18.1%; Score 633; DB 9; Length 645;  
Best Local Similarity 99.8%; Pred. No. 1.3e-152;  
Matches 644; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 2059 CTGGGAAGCCACGAGAAATTTGTAACCTGCTGGGGCGTGCACACTGTGAGGACCAATT 2118  
Db 1 CTGGGAAGCCACGAGAAATTTGTAACCTGCTGGGGCGTGCACACTGTGAGGACCAATT 60  
QY 2119 TACTTGAATTTTGAATACCTGTTGCTATGCTGATCTTCACTATCTAAGAAGTAAAGA 2178  
Db 61 TACTTGAATTTTGAATACCTGTTGCTATGCTGATCTTCACTATCTAAGAAGTAAAGA 120  
QY 2179 GAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTTCAAGTTTACCCC 2238  
Db 121 GAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTTCAAGTTTACCCC 180  
QY 2239 ACTTTCCAAATCACTCAAAATTTCCAGCTGCTGGTTCAGAGAACTTCAAGTATACCCG 2298  
Db 181 ACTTTCCAAATCACTCAAAATTTCCAGCTGCTGGTTCAGAGAACTTCAAGTATACCCG 240  
QY 2299 GACTCGGATCAATCTCAGGGCTTTCATGGAAATTTTCAAGTATACCCG 2358  
Db 241 GACTCGGATCAATCTCAGGGCTTTCATGGAAATTTTCAAGTATACCCG 300  
QY 2359 TATGAAAACCAAAAAGGCTGGAAGAAGAGGAGGACTTCAATGTGCTTACATTTGAAGAT 2418  
Db 301 TATGAAAACCAAAAAGGCTGGAAGAAGAGGAGGACTTCAATGTGCTTACATTTGAAGAT 360  
QY 2419 CTCTTTGCTTTGCATATCAAGTTCGCAAGAAATGGAATTTTCAAGTATACCCG 2478  
Db 361 CTCTTTGCTTTGCATATCAAGTTCGCAAGAAATGGAATTTTCAAGTATACCCG 420  
QY 2479 GTTCACAGAGACTGCGCCGACAGAACTGCTGTCACCCAGGAAAGTGGTGAAGATA 2538  
Db 421 GTTCACAGAGACTGCGCCGACAGAACTGCTGTCACCCAGGAAAGTGGTGAAGATA 480  
QY 2539 TGTGACTTTGGATTTGGCTGCGAGATATCATGAGTATTCACATATGTTCTCAGGGCAAT 2598  
Db 481 TGTGACTTTGGATTTGGCTGCGAGATATCATGAGTATTCACATATGTTCTCAGGGCAAT 540  
QY 2599 GCCGCTGCTGCTGAAAATGGAATGAGTGGCCCGGAAAGCCCTG-TTTTGAAGGCACTTACACCAT 2657  
Db 541 GCCGCTGCTGCTGAAAATGGAATGAGTGGCCCGGAAAGCCCTGTTTGAAGGCACTTACACCAT 600  
QY 2658 TAAGAGTGAATGCTGGTGCATATGGAATATTAATCTGTGGGAAATCTT 2702  
Db 601 TAAGAGTGAATGCTGGTGCATATGGAATATTAATCTGTGGGAAATCTT 645

RESULT 8  
US-09-796-692-8085  
;; Sequence 8085, Application US/09796692  
;; Publication No. US20020198362A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Gaiger, Alexander  
;; APPLICANT: Algate, Paul A.  
;; APPLICANT: Mannion, Jane  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER  
;; FILE REFERENCE: 2077.001200  
;; CURRENT APPLICATION NUMBER: US/09/796.692  
;; CURRENT FILING DATE: 2001-03-01  
;; PRIOR APPLICATION NUMBER: 60/186,126  
;; PRIOR FILING DATE: 2000-03-01  
;; PRIOR APPLICATION NUMBER: 60/190,479  
;; PRIOR FILING DATE: 2000-03-17  
;; PRIOR APPLICATION NUMBER: 60/200,545  
;; PRIOR FILING DATE: 2000-04-27  
;; PRIOR APPLICATION NUMBER: 60/200,303  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: 60/200,779  
;; PRIOR FILING DATE: 2000-04-28  
;; PRIOR APPLICATION NUMBER: 60/200,999  
;; PRIOR FILING DATE: 2000-05-01  
;; PRIOR APPLICATION NUMBER: 60/202,084  
;; PRIOR FILING DATE: 2000-05-04  
;; PRIOR APPLICATION NUMBER: 60/206,201  
;; PRIOR FILING DATE: 2000-05-22  
;; PRIOR APPLICATION NUMBER: 60/218,950  
;; PRIOR FILING DATE: 2000-07-14  
;; PRIOR APPLICATION NUMBER: 60/222,903  
;; PRIOR FILING DATE: 2000-08-03  
;; PRIOR APPLICATION NUMBER: 60/223,416  
;; PRIOR FILING DATE: 2000-08-04  
;; PRIOR APPLICATION NUMBER: 60/223,378  
;; PRIOR FILING DATE: 2000-08-07  
;; NUMBER OF SEQ ID NOS: 9597  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 8085  
;; LENGTH: 610  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-796-692-8085

Query Match 17.4%; Score 608.4; DB 9; Length 610;  
Best Local Similarity 99.8%; Pred. No. 2.8e-146;  
Matches 609; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2059 CTGGGAAGCCACGAGAAATTTGTAACCTGCTGGGGCGTGCACACTGTGAGGACCAATT 2118  
Db 1 CTGGGAAGCCACGAGAAATTTGTAACCTGCTGGGGCGTGCACACTGTGAGGACCAATT 60  
QY 2119 TACTTGAATTTTGAATACCTGTTGCTATGCTGATCTTCACTATCTAAGAAGTAAAGA 2178  
Db 61 TACTTGAATTTTGAATACCTGTTGCTATGCTGATCTTCACTATCTAAGAAGTAAAGA 120  
QY 2179 GAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTTCAAGTTTACCCC 2238  
Db 121 GAAAAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTTCAAGTTTACCCC 180  
QY 2239 ACTTTCCAAATCACTCAAAATTTCCAGCTGCTGGTTCAGAGAACTTCAAGTATACCCG 2298  
Db 181 ACTTTCCAAATCACTCAAAATTTCCAGCTGCTGGTTCAGAGAACTTCAAGTATACCCG 240  
QY 2299 GACTCGGATCAATCTCAGGGCTTTCATGGAAATTTTCAAGTATACCCG 2358  
Db 241 GACTCGGATCAATCTCAGGGCTTTCATGGAAATTTTCAAGTATACCCG 300  
QY 2359 TATGAAAACCAAAAAGGCTGGAAGAAGAGGAGGACTTGAATGTGCTTACATTTGAAGAT 2418  
Db 301 TATGAAAACCAAAAAGGCTGGAAGAAGAGGAGGACTTGAATGTGCTTACATTTGAAGAT 360

QY 2419 CTTCTTTGCTTGCATATCAAGTTGCCAAGGAATGGAATTTCTGAAATTTAAGTCGTGT 2478  
|||||  
Db 361 CTTCTTTGCTTGCATATCAAGTTGCCAAGGAATGGAATTTCTGAAATTTAAGTCGTGT 420  
QY 2479 GTTCACAGAGACCTGGCCGCCAGGAACGCTGTTGTACCCACCGGAAAGTGTGAAGATA 2538  
Db 421 GTTCACAGAGACCTGGCCGCCAGGAACGCTGTTGTACCCACCGGAAAGTGTGAAGATA 480  
QY 2539 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGTGATTCCAACTATGTTCTCAGGGGCAT 2598  
|||||  
Db 481 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGTGATTCCAACTATGTTCTCAGGGGCAT 540  
QY 2599 GCCCTCTGCTCTGTTAAATGGAATGTCGCCGCCAGGAACGCTGTTTGAAGGCATCTACACCAT 2658  
Db 541 GCCCTCTGCTCTGTTAAATGGAATGTCGCCGCCAGGAACGCTGTTTGAAGGCATCTACACCAT 600  
QY 2659 AAGAGTGATG 2668  
|||||  
Db 601 AAGAGTGATG 610

## RESULT 9

US-10-040-862-8085

; Sequence 8085, Application US/10040862

; Publication No. US20030078396A1

; GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander

; APPLICANT: Algate, Paul A.

; APPLICANT: Mannion, Jane

; APPLICANT: Retter, Marc

; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy

; FILE REFERENCE: Hematological Malignancies

; CURRENT APPLICATION NUMBER: US/10/040,862

; CURRENT FILING DATE: 2001-11-06

; PRIOR APPLICATION NUMBER: US 60/186,126

; PRIOR FILING DATE: 2000-03-01

; PRIOR APPLICATION NUMBER: US 60/190,479

; PRIOR FILING DATE: 2000-03-17

; PRIOR APPLICATION NUMBER: US 60/200,545

; PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: US 60/200,303

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: US 60/200,779

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: US 60/200,999

; PRIOR FILING DATE: 2000-05-01

; PRIOR APPLICATION NUMBER: US 60/202,084

; PRIOR FILING DATE: 2000-05-04

; PRIOR APPLICATION NUMBER: US 60/206,201

; PRIOR FILING DATE: 2000-05-22

; PRIOR APPLICATION NUMBER: US 60/218,950

; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: US 60/222,903

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: US 60/223,416

; PRIOR FILING DATE: 2000-08-04

; PRIOR APPLICATION NUMBER: US 60/223,378

; PRIOR FILING DATE: 2000-08-07

; PRIOR APPLICATION NUMBER: US 09/796,692

; PRIOR FILING DATE: 2001-03-01

; NUMBER OF SEQ ID NOS: 10467

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 8085

; LENGTH: 610

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-040-862-8085

## Query Match

Best Local Similarity 17.4%; Score 608.4; DB 9; Length 610;

Matches 609; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2059 CTGGGAAGCCACGAGAATATTGTGAACCTGCTGGGGCGTGCACACTGTCAAGACCAATT 2118  
Db 1 CTGGGAAGCCACGAGAATATTGTGAACCTGCTGGGGCGTGCACACTGTCAAGACCAATT 60  
QY 2119 TACTTGAATTTTGAATFACCTGTTGCTATGCTGATCTTCTCAACTATATCAAGAGTAAAGA 2178  
Db 61 TACTTGAATTTTGAATFACCTGTTGCTATGCTGATCTTCTCAACTATATCAAGAGTAAAGA 120  
QY 2179 GAAAAATTTACAGAGACTTGGACAGAGATTTCAAGGAACACAAATTTTACCC 2238  
Db 121 GAAAAATTTACAGAGACTTGGACAGAGATTTCAAGGAACACAAATTTTACCC 180  
QY 2239 ACTTTCCAATCACATCCAAATCCAGCATGCCCTGGTTCAAGAGAAGTTCAGATACACCCG 2298  
Db 181 ACTTTCCAATCACATCCAAATCCAGCATGCCCTGGTTCAAGAGAAGTTCAGATACACCCG 240  
QY 2299 GACTCGGATCAAAATCTCAGGGCTTTCATGGGAATTCATCTCACTCTGAAGATGAATGAA 2358  
Db 241 GACTCGGATCAAAATCTCAGGGCTTTCATGGGAATTCATCTCACTCTGAAGATGAATGAA 300  
QY 2359 TATGAAAAACCAAAAGGCTGGAAGAGAGAGGAGGACTTGAATGCTTACATTTGAGAT 2418  
Db 301 TATGAAAAACCAAAAGGCTGGAAGAGAGAGGAGGACTTGAATGCTTACATTTGAGAT 360  
QY 2419 CTTCTTTGCTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCGTGT 2478  
Db 361 CTTCTTTGCTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCGTGT 420  
QY 2479 GTTCACAGAGACCTGGCCGCCAGGAACGCTTGTCTCACCACCGGAAAGTGTGAAGATA 2538  
Db 421 GTTCACAGAGACCTGGCCGCCAGGAACGCTTGTCTCACCACCGGAAAGTGTGAAGATA 480  
QY 2539 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGTGATTCCAACTATGTTGTCAAGGCAAT 2598  
Db 481 TGTGACTTTGGATTGGCTCGAGATATCATGAGTGTGATTCCAACTATGTTGTCAAGGCAAT 540  
QY 2599 GCCCGTCTGCTGTAATAATGGATGGCCGCCAGGAACGCTGTTTGAAGGCATCTACACCAT 2658  
Db 541 GCCCGTCTGCTGTAATAATGGATGGCCGCCAGGAACGCTGTTTGAAGGCATCTACACCAT 600  
QY 2659 AAGAGTGATG 2668  
|||||  
Db 601 AAGAGTGATG 610

## RESULT 10

US-09-796-692-8132/c

; Sequence 8132, Application US/09796692

; Publication No. US20020198362A1

; GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander

; APPLICANT: Algate, Paul A.

; APPLICANT: Mannion, Jane

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER

; FILE REFERENCE: 2077.001200

; CURRENT APPLICATION NUMBER: US/09/796,692

; CURRENT FILING DATE: 2001-03-01

; PRIOR APPLICATION NUMBER: 60/186,126

; PRIOR FILING DATE: 2000-03-01

; PRIOR APPLICATION NUMBER: 60/190,479

; PRIOR FILING DATE: 2000-03-17

; PRIOR APPLICATION NUMBER: 60/200,545

; PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: 60/200,303

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: 60/200,779

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: 60/200,999

; PRIOR FILING DATE: 2000-05-01

; PRIOR APPLICATION NUMBER: 60/202,084

; PRIOR FILING DATE: 2000-05-04



```

; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8132
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-8132

```

Query Match	17.3%	Score 606.4	DB 9	Length 608
Best Local Similarity	99.8%	Pred. No. 9e-146		
Matches 607	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	2288	AGATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACCTCTGAAG	2347	
Db	608	AGATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACCTCTGAAG	549	
QY	2348	ATGAAATTGAATATGAAACCAAAAAAGCTTGGAAAGAGAGGAGACTTGAATGCTGCTTA	2407	
Db	548	ATGAAATTGAATATGAAACCAAAAAAGCTTGGAAAGAGAGGAGACTTGAATGCTGCTTA	489	
QY	2408	CATTGGAAGATCTCTTTGCTTTGCATATCAAGTTGCCAAGGAATGGAAATTTCTGGAAT	2467	
Db	488	CATTGGAAGATCTCTTTGCTTTGCATATCAAGTTGCCAAGGAATGGAAATTTCTGGAAT	429	
QY	2468	TTAAGTCGTGTTTCACAGAGACTGCCGCCAGGAACGTGCTTGTCACCCACGGGAAG	2527	
Db	428	TTAAGTCGTGTTTCACAGAGACTGCCGCCAGGAACGTGCTTGTCACCCACGGGAAG	369	
QY	2528	TGTTGAAGATATGTGACTTTTGATTTGGATTGCTCGAGATATCATGAGTGATTCACACTATGTTG	2587	
Db	368	TGTTGAAGATATGTGACTTTTGATTTGGATTGCTCGAGATATCATGAGTGATTCACACTATGTTG	309	
QY	2588	TCAGGGGCAATGCCCGTCTGCCCTGTAATAATGGATGGCCCCGAAAGCCTGTTGAAGGCA	2647	
Db	308	TCAGGGGCAATGCCCGTCTGCCCTGTAATAATGGATGGCCCCGAAAGCCTGTTGAAGGCA	249	
QY	2648	TCACACCAATTAAGAGTGATGCTGGTGCATATGGAATATTACTGTGGGAATCTTCTCAC	2707	
Db	248	TCACACCAATTAAGAGTGATGCTGGTGCATATGGAATATTACTGTGGGAATCTTCTCAC	189	
QY	2708	TTGGTGTGAATCCTTACCCTGGCAATCCGGTTGATGCTACTTCTACAAACTGATTCAAA	2767	
Db	188	TTGGTGTGAATCCTTACCCTGGCAATCCGGTTGATGCTACTTCTACAAACTGATTCAAA	129	
QY	2768	ATGGATTAAATATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATATGCAAT	2827	
Db	128	ATGGATTAAATATGGATCAGCCATTTTATGCTACAGAAGAAATATACATTATATGCAAT	69	
QY	2828	CTGCTGGGCTTTTGACTCAAGGAACGGGCATCTCCCTAAATTTGACCTTCGTTTTTAG	2887	
Db	68	CTGCTGGGCTTTTGACTCAAGGAACGGGCATCTCCCTAAATTTGACCTTCGTTTTTAG	9	
QY	2888	GATGTCAG 2895		
Db	8	GATGTCAG 1		

RESULT 11  
US-10-040-862-8132/C  
; Sequence 8132, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander

```

: APPLICANT: Algate, Paul A.
: APPLICANT: Mannion, Jane
: APPLICANT: Retter, Marc
: APPLICANT: Corixa Corporation
: TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther
: FILE OF INVENTION: Hematological Malignancies
: FILE REFERENCE: 014058-013520US
: CURRENT APPLICATION NUMBER: US/10/040,862
: CURRENT FILING DATE: 2001-11-06
: PRIOR APPLICATION NUMBER: US 60/186,126
: PRIOR FILING DATE: 2000-03-01
: PRIOR APPLICATION NUMBER: US 60/190,479
: PRIOR FILING DATE: 2000-03-17
: PRIOR APPLICATION NUMBER: US 60/200,545
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: US 60/200,303
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: US 60/200,779
: PRIOR FILING DATE: 2000-04-28
: PRIOR APPLICATION NUMBER: US 60/200,999
: PRIOR FILING DATE: 2000-05-01
: PRIOR APPLICATION NUMBER: US 60/202,084
: PRIOR FILING DATE: 2000-05-04
: PRIOR APPLICATION NUMBER: US 60/206,201
: PRIOR FILING DATE: 2000-05-22
: PRIOR APPLICATION NUMBER: US 60/218,950
: PRIOR FILING DATE: 2000-07-14
: PRIOR APPLICATION NUMBER: US 60/222,903
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: US 60/223,416
: PRIOR FILING DATE: 2000-08-04
: PRIOR APPLICATION NUMBER: US 60/223,378
: PRIOR FILING DATE: 2000-08-07
: PRIOR APPLICATION NUMBER: US 09/796,692
: PRIOR FILING DATE: 2001-03-01
: NUMBER OF SEQ ID NOS: 10467
: SOFTWARE: FastSEQ for Windows Version 3.0
: SEQ ID NO 8132
: LENGTH: 608
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-040-862-8132

```

	Query Match	17.3%	Score 606.4	DB 9	Length 608	
	Best Local Similarity	99.8%	Pred. No. 9e-146			
	Matches 607	Conservative 0	Mismatches 1	Indels 0	Gaps 0	
QY	2288	AGATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAG	2347			
Db						
QY	608	AGATACACCGGACTCGGATCAAAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAG	549			
Db						
QY	2348	ATGAAATTGAATATGAAACCAAAAAGGCTTGAAGAAGAGGAGGACTTCAAATGTGCTTA	2407			
Db						
QY	548	ATGAAATTGAATATGAAACCAAAAAGGCTTGAAGAAGAGGAGGACTTCAAATGTGCTTA	489			
Db						
QY	2408	CATTGAAGATCTCTTTGCTTTGCATATCAAGTTGCCAAAGGAATTTCTTGGAAAT	2467			
Db						
QY	488	CATTGAAGATCTCTTTGCTTTGCATATCAAGTTGCCAAAGGAATTTCTTGGAAAT	429			
Db						
QY	2468	TTAAGTCGTGTGTTACAGAGACCTGGCCCGCAGGAACGTGTGTACCCAGGGAAAG	2527			
Db						
QY	428	TTAAGTCGTGTGTTACAGAGACCTGGCCCGCAGGAACGTGTGTACCCAGGGAAAG	369			
Db						
QY	2528	TGTTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGACTGATTCCAACTATGTTG	2587			
Db						
QY	368	TGTTGAAGATATGTGACTTTGGATTGGCTCGAGATATCATGACTGATTCCAACTATGTTG	309			
Db						
QY	2588	TCAGGGSCAATGCCCGCTGCTCTGAAAATGGATGGCCCCGAAAGCCTGTTTGAAGGCA	2647			
Db						
QY	308	TCAGGGSCAATGCCCGCTGCTCTGAAAATGGATGGCCCCGAAAGCCTGTTTGAAGGCA	249			
Db						
QY	2648	TCTACACCATTAAGAGTGATGCTGTGTCATATGGAATATTACTGTGGAAATCTTCTCAC	2707			
Db						

Db 248 TCTACACCATTAAGAGTGATGCTGTGCATATGGAATATTACTGTGGAAATCTTCTCAC 189  
QY 2708 TTGGTGTGAATCTTACCTGGGCAATCCGGTTGATGCTAACTTCTCAAACTGATTCAAA 2767  
Db 188 TTGGTGTGAATCTTACCTGGGCAATCCGGTTGATGCTAACTTCTCAAACTGATTCAAA 129  
QY 2768 ATGGATTTAAATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2827  
Db 128 ATGGATTTAAATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 69  
QY 2828 CTTGCTGGGCTTTTGAATCAAGGAAGGCGCATCTTCCCTAAATTTGACTTCGTTTATAG 2887  
Db 68 CTTGCTGGGCTTTTGAATCAAGGAAGGCGCATCTTCCCTAAATTTGACTTCGTTTATAG 9  
QY 2888 GATGTCAG 2895  
Db 8 GATGTCAG 1

RESULT 12  
US-09-796-692-8927  
; Sequence 8927, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09796,692  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; NUMBER OF SEQ ID NOS: 9597  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8927  
; LENGTH: 605  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-796-692-8927  
Query Match 17.3%; Score 605; DB 9; Length 605;  
Best Local Similarity 100.0%; Pred. No. 2.1e-145;  
Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2059 CTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGCGGTGCACACTGTCCAGGACCAAT 2118  
Db 1 CTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGCGGTGCACACTGTCCAGGACCAAT 60  
QY 2119 TACTTGATTTTGAATACTGTTGCTATGGTGATCTTCTCAACTATCTAAGAAGTAAAGA 2178

Db 61 TACTTGATTTTGAATACTGTTGCTATGGTGATCTTCTCAACTATCTAAGAAGTAAAGA 120  
QY 2179 GAAAAATTTTCAGAGACTTGGACAGAGATTTTCAAGGAACACAAATTCAGTTTTTACCCC 2238  
Db 121 GAAAAATTTTCAGAGACTTGGACAGAGATTTTCAAGGAACACAAATTCAGTTTTTACCCC 180  
QY 2239 ACTTTCCAAATCACAATCCAAATTTCCAGCATGCTGTTTCAAGGAGAGTTTCAGATACACCG 2298  
Db 181 ACTTTCCAAATCACAATCCAAATTTCCAGCATGCTGTTTCAAGGAGAGTTTCAGATACACCG 240  
QY 2299 GACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTTCACTCTGAAGATGAAATTTGAA 2358  
Db 241 GACTCGGATCAAAATCTCAGGCTTTCATGGGAATTCATTTCACTCTGAAGATGAAATTTGAA 300  
QY 2359 TATGAAAAACCAAAAGGCTGGAAGAGGAGGACTTGAATGCTTACATTTGAAGAT 2418  
Db 301 TATGAAAAACCAAAAGGCTGGAAGAGGAGGACTTGAATGCTTACATTTGAAGAT 360  
QY 2419 CTTCTTGTGCTTGCATATCAAGTTGCCAAGGAATGGAATTTCTGGAATTTAAGTCGTGT 2478  
Db 361 CTTCTTGTGCTTGCATATCAAGTTGCCAAGGAATGGAATTTCTGGAATTTAAGTCGTGT 420  
QY 2479 GTTCACAGAGACTGCGCCGAGGAGGAGGACTTGAATGCTTGAAGATGAAATTTGAAGAT 2538  
Db 421 GTTCACAGAGACTGCGCCGAGGAGGAGGACTTGAATGCTTGAAGATGAAATTTGAAGAT 480  
QY 2539 TGTGACTTTGGATTTGGCTGCGAGATATCATGAGTGAATTCACACTATGTTGTCAGGGGCAAT 2598  
Db 481 TGTGACTTTGGATTTGGCTGCGAGATATCATGAGTGAATTCACACTATGTTGTCAGGGGCAAT 540  
QY 2599 GCCCGTCTGCTGTAATAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 2658  
Db 541 GCCCGTCTGCTGTAATAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 600  
QY 2659 AAGAG 2663  
Db 601 AAGAG 605

RESULT 13  
US-10-040-862-8927  
; Sequence 8927, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Ther  
; TITLE OF INVENTION: Hematological Malignancies  
; FILE REFERENCE: 014058-013520US  
; CURRENT APPLICATION NUMBER: US/10/040,862  
; CURRENT FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: US 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: US 09/796,692  
; PRIOR FILING DATE: 2001-03-01  
; NUMBER OF SEQ ID NOS: 10467  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8927  
; LENGTH: 605  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-040-862-8927

Query Match 17.3%; Score 605; DB 9; Length 605;  
Best Local Similarity 100.0%; Pred. No. 2.1e-145;  
Matches 605; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2059 CTGGGAAGCCACGAGATATTGTGAACCTGCTGGGGCGTGCACACTGTCAAGACCAATT 2118  
Db 1 CTGGGAAGCCACGAGATATTGTGAACCTGCTGGGGCGTGCACACTGTCAAGACCAATT 60  
QY 2119 TACTTGATTTTGAATCTGTCTGATGATCTTCTCAACTATCTAAGAAGTAAAGA 2178  
Db 61 TACTTGATTTTGAATCTGTCTGATGATCTTCTCAACTATCTAAGAAGTAAAGA 120  
QY 2179 GAAAAATTCACAGGACTTGACAGAGATTTTCAAGGAACAATTTTCAAGTTTACCCC 2238  
Db 121 GAAAAATTCACAGGACTTGACAGAGATTTTCAAGGAACAATTTTCAAGTTTACCCC 180  
QY 2239 ACTTTCCAATCACATCCAAATTCAGCATGCTGTTCAAGAGAAGTTTCAGATACACCCG 2298  
Db 181 ACTTTCCAATCACATCCAAATTCAGCATGCTGTTCAAGAGAAGTTTCAGATACACCCG 240  
QY 2299 GACTCGGATCAAAATCTCAGGCTTCATCGGAATTCATTTCACTCTGAAGATGAATGAA 2358  
Db 241 GACTCGGATCAAAATCTCAGGCTTCATCGGAATTCATTTCACTCTGAAGATGAATGAA 300  
QY 2359 TATGAAACCAAAAAGGCTGGAAGAAGAGAGAGACATTTGAATGCTTACATTTGAAGAT 2418  
Db 301 TATGAAACCAAAAAGGCTGGAAGAAGAGAGAGACATTTGAATGCTTACATTTGAAGAT 360  
QY 2419 CTCTTTGCTTTGCATATCAATGTCGCAAGAGATGGAATTTCTGGAAATTTAAGTCGTGT 2478  
Db 361 CTCTTTGCTTTGCATATCAATGTCGCAAGAGATGGAATTTCTGGAAATTTAAGTCGTGT 420  
QY 2479 GTTCACAGAGACCTGGCCGCGCAGGAAGCTGTTGTCAACCCACGGGAAAGTGGTGAAGATA 2538  
Db 421 GTTCACAGAGACCTGGCCGCGCAGGAAGCTGTTGTCAACCCACGGGAAAGTGGTGAAGATA 480  
QY 2539 TGTGACTTTGGATTGGCTCGAGATATCATGATGATTCCAACTATGTTGTCAGGGGCAAT 2598  
Db 481 TGTGACTTTGGATTGGCTCGAGATATCATGATGATTCCAACTATGTTGTCAGGGGCAAT 540  
QY 2599 GCCCGTCTGCTGTAATATGGATGGCCCCCAGAGCTGTTTGAAGGCATCTACACCAAT 2658  
Db 541 GCCCGTCTGCTGTAATATGGATGGCCCCCAGAGCTGTTTGAAGGCATCTACACCAAT 600  
QY 2659 AAGAG 2663  
Db 601 AAGAG 605

RESULT 14  
US-09-796-692-8273  
; Sequence 8273, Application US/09796692  
; Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THER  
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09/796,692  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; NUMBER OF SEQ ID NOS: 9597  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8273  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: unsure  
; LOCATION: (586)  
; OTHER INFORMATION: n=A,T,C or G  
US-09-796-692-8273

Query Match 17.1%; Score 598.4; DB 9; Length 601;  
Best Local Similarity 99.7%; Pred. No. 1e-143;  
Matches 599; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2059 CTGGGAAGCCACGAGATATTGTGAACCTGCTGGGGCGTGCACACTGTCAAGACCAATT 2118  
Db 1 CTGGGAAGCCACGAGATATTGTGAACCTGCTGGGGCGTGCACACTGTCAAGACCAATT 60  
QY 2119 TACTTGATTTTGAATCTGTCTGATGATCTTCTCAACTATCTAAGAAGTAAAGA 2178  
Db 61 TACTTGATTTTGAATCTGTCTGATGATCTTCTCAACTATCTAAGAAGTAAAGA 120  
QY 2179 GAAAAATTCACAGACTTGACAGAGATTTTCAAGGAACAATTTTCAAGTTTACCCC 2238  
Db 121 GAAAAATTCACAGACTTGACAGAGATTTTCAAGGAACAATTTTCAAGTTTACCCC 180  
QY 2239 ACTTTCCAATCACATCCAAATTCAGCATGCTGTTCAAGAGAAGTTTCAGATACACCCG 2298  
Db 181 ACTTTCCAATCACATCCAAATTCAGCATGCTGTTCAAGAGAAGTTTCAGATACACCCG 240  
QY 2299 GACTCGGATCAAAATCTCAGGCTTCATCGGAATTCATTTCACTCTGAAGATGAATGAA 2358  
Db 241 GACTCGGATCAAAATCTCAGGCTTCATCGGAATTCATTTCACTCTGAAGATGAATGAA 300  
QY 2359 TATGAAACCAAAAAGGCTGGAAGAAGAGAGAGACATTTGAATGCTTACATTTGAAGAT 2418  
Db 301 TATGAAACCAAAAAGGCTGGAAGAAGAGAGAGACATTTGAATGCTTACATTTGAAGAT 360  
QY 2419 CTCTTTGCTTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCGTGT 2478  
Db 361 CTCTTTGCTTTGCATATCAAGTTGCCAAAGGAATGGAATTTCTGGAATTTAAGTCGTGT 420

2059 CTGGAGCCACGGAATATTGTGAACCTGCTGGGGCGTGCACACTGTACAGACCAATT 2118  
1 CTGGAGCCACGGAATATTGTGAACCTGCTGGGGCGTGCACACTGTACAGACCAATT 60  
2119 TACTTGATTTTGAATACTGTGTATGTGTATCTTCAACTATCTAAGAAGTAAAGA 2178  
61 TACTTGATTTTGAATACTGTGTATGTGTATCTTCAACTATCTAAGAAGTAAAGA 120  
2179 GAAAAATTTTCAAGAGACTTTTCAAGAGACACAATTTTCAAGAGTAAAGTAAAGA 2238  
121 GAAAAATTTTCAAGAGACTTTTCAAGAGACACAATTTTCAAGAGTAAAGTAAAGA 180  
2239 ACTTTTCAATCATCTCAAAATTCAGATGCTGCTTCAAGAGAACTTTCAGATACACCCG 2298  
181 ACTTTTCAATCATCTCAAAATTCAGATGCTGCTTCAAGAGAACTTTCAGATACACCCG 240  
2299 GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTTTCACTCTTGAAGTAAATGAA 2358  
241 GACTCGGATCAAAATCTCAGGGCTTCATGGGAATTTTCACTCTTGAAGTAAATGAA 300  
2359 TATGAAACCAAAAAAGGCTGGAAGAGAGAGGAGCTTGAATGTCTTACATTGAGAT 2418  
301 TATGAAACCAAAAAAGGCTGGAAGAGAGAGGAGCTTGAATGTCTTACATTGAGAT 360  
2419 CTCTTTTGTCTTGCATATCAAGTTGCAAGAGAAATGGAATTTTCTGGAATTTAAGTCGTG 2478  
361 CTCTTTTGTCTTGCATATCAAGTTGCAAGAGAAATGGAATTTTCTGGAATTTAAGTCGTG 420  
2479 GTTCACAGAGACCTGCGCCGACGAGACGCTGCTTGTACCCACCGGAAAGTGGTGAAGATA 2538  
421 GTTCACAGAGACCTGCGCCGACGAGACGCTGCTTGTACCCACCGGAAAGTGGTGAAGATA 480  
2539 TGTGACTTTTGGATTGGCTCGAGATATCATGAGTGATTTCCAACTATGTTGTCAGGGGCAAT 2598  
481 TGTGACTTTTGGATTGGCTCGAGATATCATGAGTGATTTCCAACTATGTTGTCAGGGGCAAT 540  
2599 GCCCGTCTGCTGTAAATGGATGGCCCGGAAAGCCTGTTTGAAGGCATCTACACCAT 2658  
541 GCCCGTCTGCTGTAAATGGATGGCCCGGAAAGCCTGTTTGAAGGCATCTACACCAT 600  
2659 A 2659  
601 A 601

Search completed: May 25, 2003, 17:39:07  
Job time : 406.768 secs

2479 GTTCACAGAGACCTGCGCCGACGAGAGCTGCTTCTACCCACGGAAGAGTGGTGAAGATA 2538  
421 GTTCACAGAGACCTGCGCCGACGAGAGCTGCTTCTACCCACGGAAGAGTGGTGAAGATA 480  
2539 TGTGACTTTTGGATTGGCTCGAGATATCATGAGTGATTTCCAACTATGTTGTCAGGGGCAAT 2598  
481 TGTGACTTTTGGATTGGCTCGAGATATCATGAGTGATTTCCAACTATGTTGTCAGGGGCAAT 540  
2599 GCCCGTCTGCTGTAAATGGATGGCCCGGAAAGCCTGTTTGAAGGCATCTACACCAT 2658  
541 GCCCGTCTGCTGTAAATGGATGGCCCGGAAAGCCTGTTTGAAGGCATCTACACCAT 600  
2659 A 2659  
601 A 601

RESULT 15  
US-10-040-862-8273  
; Sequence 8273, Application US/10040862  
; Publication No. US20030078396A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; TITLE OF INVENTION: Hematological Malignancies  
; FILE REFERENCE: 014058-013520US  
; CURRENT APPLICATION NUMBER: US/10/040,862  
; CURRENT FILING DATE: 2001-11-06  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: US 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: US 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: US 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: US 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: US 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: US 09/796,692  
; PRIOR FILING DATE: 2001-03-01  
; NUMBER OF SEQ ID NOS: 10467  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8273  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (586)  
; OTHER INFORMATION: n=A,T,C or G

Query Match 17.1%; Score 598.4; DB 9; Length 601;  
Best Local Similarity 99.7%; Pred. No. 1e-143;  
Matches 599; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

US-10-040-862-8273

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 25, 2003, 03:00:43 ; Search time 4597.52 Seconds

(without alignments)

12332.831 Million cell updates/sec

Title: US-09-919-408-3

Perfect score: 3501

Sequence: 1 CGAGGCGCATCCGAGGCT.....AAAAAAAAAAAAAAAAAAAA 3501

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154056 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estnu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_estl:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: gb\_gss:\*

18: em\_gss\_hum:\*

19: em\_gss\_inv:\*

20: em\_gss\_pln:\*

21: em\_gss\_vrt:\*

22: em\_gss\_fun:\*

23: em\_gss\_mam:\*

24: em\_gss\_mus:\*

25: em\_gss\_other:\*

26: em\_gss\_pro:\*

27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	650.6	18.6	714	13	BI461248
2	509.6	14.6	621	13	BI360262
3	486.2	13.9	765	9	AI323253
4	440	12.6	604	12	BF523018
5	437.8	12.5	460	10	AV713950
6	402.6	11.5	618	9	AA120050

7	400.4	11.4	455	13	BM147250
8	389.6	11.1	507	13	BM484050
9	340.4	9.7	3312	11	AK004947
10	325.6	9.3	445	10	BE651447
11	318.6	9.1	5174	11	BC026713
12	314.4	9.0	522	10	BB644407
13	305.4	8.7	326	13	BM144393
14	299.6	8.6	646	10	AW772610
15	283.4	8.1	659	10	AW057705
16	274.8	7.8	588	10	BE679272
17	249.2	7.1	515	14	BQ556177
18	240.4	6.9	394	12	BG145808
19	236.6	6.8	660	10	BE047675
20	230.4	6.6	276	12	BF176247
21	228.4	6.5	739	12	BF159205
22	228	6.5	480	9	AI150354
23	223.8	6.4	666	13	BG934591
24	218.6	6.2	346	13	BM253635
25	218.6	6.2	346	13	BM253796
26	218.2	6.2	503	13	BI481351
27	216.4	6.2	832	13	BI694328
28	214	6.1	357	9	AI556371
29	213.6	6.1	794	13	BI111563
30	212.8	6.1	347	13	BM253661
31	212.2	6.1	407	10	AV596131
32	210.4	6.0	940	9	AI325368
33	209.6	6.0	747	9	AI323643
34	204.8	5.8	702	12	BF163671
35	204.8	5.8	848	13	BI107537
36	202.6	5.8	669	13	BJ012204
37	201.8	5.8	531	12	BF320739
38	201.6	5.8	666	9	AL047744
39	200.8	5.7	983	12	BF782576
40	197.8	5.6	201	10	AW059681
41	197.4	5.6	412	12	BE853746
42	196.6	5.6	461	9	AI092636
43	194.8	5.6	575	12	BF930052
44	194.4	5.6	696	14	BM951127
45	193.8	5.5	665	13	BG922846

ALIGNMENTS

RESULT 1	BI461248	BI461248	714 bp	mRNA	linear	EST 21-AUG-2001				
LOCUS	603206574F1	NIH_MGC_97	Homo sapiens	CDNA clone	IMAGE:5272266 5',					
DEFINITION	mRNA sequence.									
ACCESSION	BI461248	GI:15251904								
VERSION	BI461248.1	EST.								
KEYWORDS	human.									
SOURCE	human.									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.									
AUTHORS	NIH-MGC	http://mgc.nci.nih.gov/.								
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)									
JOURNAL	Unpublished (1999)									
COMMENT	Contact: Robert Strausberg, Ph.D.									
	Email: cgabbs-remail.nih.gov									
	Tissue Procurement: Miklos Palkovits, M.D., Ph.D.									
	CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshiyuki and Piero Carninci (RIKEN)									
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)									
	DNA Sequencing by: Incyte Genomics, Inc.									
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:									
	http://image.llnl.gov									
	Plate: LLAM11687	row: c	column: 19							
	High quality sequence stop: 709.									
	Location/Qualifiers									

FEATURES

```

source
1. .714
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5272266"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-xhoI (gtcagag
); Oligo-dr primed using primer 5'-tttttttttttttttttVN-3',
size-selected for average insert size 2.2 kb and
normalized to R0F 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT      182 a  171 c  191 g  170 t
ORIGIN

Query Match      18.68; Score 650.6; DB 13; Length 714;
Best Local Similarity 97.8%; Pred. No. 8.3e-100;
Matches 670; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

Qy 1 CGAGGCGGCATCGAGGCTGGGCGCGCGC-CCTGGGGACCCCGGGCTCGGAGGCCAT 59
Db 30 CCAGGCGGCATCGAGGCTGGGCGCGCGCGCTGGGACCCCGGGCTCGGAGGCCAT 89

Qy 60 GCCGGCGTTGGCGCGACGGCGGACCGTGCCTGCTGCTGTTTTTCTGCAATGAT 119
Db 90 GCCGGCGTTGGCGCGACGGCGGCGAGCTGCCGCTGCTGTTTTTCTGCAATGAT 149

Qy 120 ATTTGGGACATTAACAATCAAGATCGCCTGTCATCAAGTGTCTTTAATCAATCAATA 179
Db 150 ATTTGGGACATTAACAATCAAGATCGCCTGTCATCAAGTGTCTTTAATCAATCAATA 209

Qy 180 GAACAATGATTCACAGTGGGGAAGTCATCATCATATATCCCATGGTATCAGAATCCCCGGA 239
Db 210 GAACAATGATTCACAGTGGGGAAGTCATCATCATATATCCCATGGTATCAGAATCCCCGGA 269

Qy 240 AGACCTCGGGTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCGCTGT 299
Db 270 AGACCTCGGGTGCGTTGAGACCCAGAGCTCAGGACAGTGTACGAAGCTGCCGCTGT 329

Qy 300 GGAAGTGGATGTATCTGTTCCATCACACTGCAAGTGTGTCGATGCCCGCAGGGAACAT 359
Db 330 GGAAGTGGATGTATCTGTTCCATCACACTGCAAGTGTGTCGATGCCCGCAGGGAACAT 389

Qy 360 TTCCTGTCTCTGGGTCTTTAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGTATTACA 419
Db 390 TTCCTGTCTCTGGGTCTTTAAGCACAGCTCCCTGAAATGCCAGCCACATTTTGTATTACA 449

Qy 420 AAACAGAGGAGCTGTGTTCCATGGTCAATTTGAAATGACAGAAACCCAAAGCTGGAGAATA 479
Db 450 AAACAGAGGAGTAGTTTCCATGGTCAATTTGAAATGACAGAAACCCAAAGCTGGAGAATA 509

Qy 480 CCTACTTTTATTACAGAGTGAAGCTACCAATATACACAATATGTTTACAGTGAATTAAG 539
Db 510 CCTACTTTTATTACAGAGTGAAGCTACCAATATACACAATATGTTTACAGTGAATTAAG 569

Qy 540 AAATACCTTGCTTTACACATTAAGAGACCTTACTTTAGAAAAATGGAACACGAGCGC 599
Db 570 AAATACCTTGCTTTACACATTAAGAGACCTTACTTTAGAAAAATGGAACACGAGCGC 629

Qy 600 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTGGGATTC 659
Db 630 CTGCTCTGCATATCTGAGAGCGTTCCAGAGCCGATCGTGAATGGGTGCTTTGGGATTC 689

Qy 660 ACAGGGGAAAGCTGTAAGAGAA 684
Db 690 ACAGGGGAAAGCTGTAAGAGCA 714

```

LOCUS	BI360262	621 bp	mrna	linear	EST 01-AUG-2000
DEFINITION	387099 MARC 2PTG Sus scrofa cDNA 5', mRNA sequence.				
ACCESSION	BI360262				
VERSION	BI360262.1 GI:15056290				
KEYWORDS	EST.				
SOURCE	pig.				
ORGANISM	Sus scrofa				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.				
AUTHORS	1 (bases 1 to 621) Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E., Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W. and Keele, J.W.				
TITLE	Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Smith.TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing. Bases called and alt.trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.				
FEATURES	PCR Primers FORWARD: AGGAAACAGCTATGACCAT BACKWARD: GTTTCGCCAGTCAGCAGC Plate: 139 row: H column: 2 Seq primer: ATTTAGGTGACACTATAG. Location/Qualifiers 1..621 /organism="Sus scrofa" /db_xref="taxon:9823" /clone_lib="MARC 2PTG" /tissue_type="pooled" /lab_host="DH10B" /note="vector: pCMV SPORT6; Site.1: NotI; Site.2: SalI; Library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."				
BASE COUNT	190 a 132 c 141 g 158 t				
ORIGIN					
Query Match	14.6%; Score 509.6; DB 13; Length 621;				
Best Local Similarity	88.9%; Pred. No. 4.4e-76;				
Matches	551; Conservative 0; Mismatches 69; Indels 0; Gaps 0;				
Qy	1725	CGTTTTAACCTGCTAATTGTCACAGTACAAAAGCAATTTAGGTATGAAGCCAGCT	1784		
Db	2	CGTTTTAACCATGCTAATTTGTCACAGTACAAAAGCGATTCCGGTACGAAGCCAGCT	61		
Qy	1785	ACAGATGTCACAGGTGACCGGCTCCTCAGATAAATGAGTACTTCTACGGTTGATTTTCAGAGA	1844		
Db	62	GCAGATGGTCGAGTGACGGGGTCCCTGGATACGACTCTTCACATCGACTTCAGAGA	121		
Qy	1845	ATATGAATATGATCTCAAAATGGGAGTTTCCAAGAGAAAATTTAGAGTTTGGGAAGGTACT	1904		
Db	122	ATATGAATATGACCTCAAAATGGGAGTTTCCCGAGGAAAATTTAGAAATTTGGGAAGGTCT	181		
Qy	1905	AGGATCAGGTGCTTTTGGAAAGTGATGAACGCAACAGCTTATGGAATTTACCAAAACAGG	1966		
Db	182	GGGATCCGGCGCTTTTGGAAAGTGATGAACCAACTGCCCTATGGAATTCAGTAAACTGG	241		
Qy	1965	AGTCTCAATCCAGGTGCCGTCAAAATGCTCAAGAGAAAACACAGCTCTCGAAAGAG	2026		
Db	242	AGTGTCAATCCAGGTGGCAGTCAAAATGCTGAAAGAAAGCAGCAGCTCGGACACGAGA	301		
Qy	2025	GGCACTCATGTGCAAGACTCAAGATGATGACCCAGCTGGGAAGCCACAGAGATATTGTGAA	2088		
Db	302	GGCTCTCATGTCTGAACTCAAAATGATGACCCACCTGGCAGCCACAGAGATATAGTGA	361		
Qy	2085	CTGCTGGGGCGGTGCACACTGTGCAGGACCAATTTACTTGAATTTTGAATCTACTGTGCTA	2144		

RESULT 2  
BI360262

Db 362 CTGCTGGCGCCTGCACCCCTGTCAGGCCCATTTACTTGATTTTGAATATTGTGCTA 421  
 QY 2145 TGGTGATCTTCTCAACTATCTAAGAAGTAAAGAGAGAAAATTTTCAGAGACTTGGACAGA 2204  
 Db 422 CGGTGATCTTCTCAACTATCTAAGAAGTAAAGAGAGAAAATTTTCATAGGACATGGACCGA 481  
 QY 2205 GATTTTCAGGAGACAAATTCAGTTTATACCCCTTTCCCAATCCATCCAAATTCAG 2264  
 Db 482 GATTTTCAGGAGACAAATTTTCAGTTTATACCCCTTTCCCAATCCATCCAAATTCAG 541  
 QY 2265 CATGCTGCTTCAAGAGAGTTTCAGATACACCCGAGCTCGGATCAAAATCTCAGGGCTTCA 2324  
 Db 542 TATGCCGGTTCAAGAGAGTTCAATACACCCGAGCTCGGATCCATCTCAGGATTCAA 601  
 QY 2325 TGGGAATTCATTTCACTCTG 2344  
 Db 602 TGGGAATTCATTTCACTCTG 621

## RESULT 3

AI323253

LOCUS

DEFINITION

mp8b02.y1 Soares\_thymus\_2NDMT Mus musculus cDNA clone IMAGE:576267  
 5' similar to gb:U02687 FL CYTOKINE RECEPTOR PRECURSOR (HUMAN);  
 gb:X59398 Mouse Flt3 mRNA for tyrosine kinase receptor of the PDGF  
 (MOUSE);, mRNA sequence.

ACCESSION

AI323253

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 765)  
 Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
 Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
 Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
 Waterston, R.

TITLE The WashU-HHMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:350915

This read is a RESEQUENCE of a previously sequenced mouse clone

correct orientation)

Seq primer: -40RP from Gibco

High quality sequence stop: 423.

Location/Qualifiers

## FEATURES

source

1..765

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone\_image="IMAGE:576267"

/sex="male"

/tissue\_type="thymus"

/dev\_stage="4 weeks"

/lab\_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA  
 was primed with a Not I - oligo(dT) primer [5',  
 TGTTACCATCTGAGTGGGAGCGCCGGTGTGTTTTTTTTTTTTTTTTTTT  
 3']; double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT73 vector. RNA  
 provided by Dr. Bertrand Jordan. Library went through two  
 rounds of normalization, and was constructed by Bento  
 Soares and M. Fatima Bonaldo."  
 BASE COUNT 203 a 176 c 182 g 200 t 4 others  
 ORIGIN

Query Match 13.9%; Score 486.2; DB 9; Length 765;

Best Local Similarity 81.5%; Pred. No. 3.5e-72;

Matches 574; Conservative 0; Mismatches 129; Indels 1; Gaps 1;

QY 1582 GGGTTCCTCGTCAAGTGTGTGCATACAAATTCCTTGGCACATCTTGTGAGACGATCCTT 1641  
 Db 1 GGGCTTCTCGTCAAAATGCTGTGGCTACAAATCTATGGCACGCTCTTGGCAACCATCTTT 60

QY 1642 TTAACACTCTCCAGGCCCTTCCCTTTCATCCACAGACAACATCTCATTTCTATGCAACAATT 1701  
 Db 61 TTAACACTCACCAGGCCCTTCCCTTTCATCCACAGACAACATCTCTCTTATGGACCATT 120

QY 1702 GGTGTTTGTCTCTCTCTTCATTTGCTTTTAAACCTCTGTAATTTGTTCACAAGTACAAAAG 1761  
 Db 121 GGGCTCTGTCTCCCTTTCATTTGTTGTCATTTGTTGATCTGCCACAATACAAAAG 180

QY 1762 CAATTTAGTGTATCAAAAGCCAGCTACAGATGGTACAGGTGACCGCTCCTCAGATAAATGAG 1821  
 Db 181 CAATTTAGTACAGAGTACGTGCGAGATGATCCAGTGACTGGCCCCCTTGGATAACGAG 240

QY 1822 TACTTCTAGCTTGATTTCAGAGAAATATGAATATGATCTCAATATGGGAGTTTCCAAAGAA 1881  
 Db 241 TACTTCTAGCTTGACTTCAGGGACTATGAATATGACCTTAAGTGGGAGTTCCCGAGAGAG 300

QY 1882 AATTTAGATTTGGGAGGTACTAGGATCAGTGGTCTTTTGGAAAAGTGTAGACGCAACA 1941  
 Db 301 AACTTAGAGTTTGGGAGGTCTTGGGCTCTGGGCTTTCGGGAGGTGATGAACGCCACG 360

QY 1942 GCTTATGGATTAAGCAAAACAGAGTCTCAATCAGGTTGCCGTCAAAATGCTCAAGAA 2001  
 Db 361 GCTTATGGATTAAGTAAACGGGAGTCTCAATCAGGTTGGGCTGGAAGATGCTTAAAGAG 420

QY 2002 AAAGCAGACAGCTCTCTGAAAGAGAGGACCTCATGTCAGAATCAAGATGATGACCCAGCTG 2061  
 Db 421 AAAGCTGACAGCTGTGAAAGAGAGGCTCTCATGTCGAGCTCAAAATGATGACCCACCTT 480

QY 2062 GGAAGCCACAGAAATATTTGAAACCTGCTGGGCGGTGACACTGTCTCAGGACCAATTTAC 2121  
 Db 481 GGACAACATGACCAACATCGTGGATTTGTTGGGCTTCCACACTGCCAGGCCAGTGCAC 540

QY 2122 TTGATTTTCAATACTGTTGCTATGGTGTCTTCAACTATCTAAGAAGTAAAGAGAA 2181  
 Db 541 TGGATTTTGAATATGGTGTCTATGGAGACCTCTCAACTACCTAAGAAGTAAAGAGAG 600

QY 2182 AAATTTTCAAGGACTTGGACAGAGATTTTCAAGGAACACAAATTTACGTTTTTACCCCACT 2241  
 Db 601 AATGTTTCAAGGACATGGACAGAGATTTTAAAGACCATAA-TTCAGTTTTTACCTTACT 659

QY 2242 TTCCAATCATCCAAATTCACAGATGCTGCTGGTTTCAAGAGAAAT 2285  
 Db 660 TTCCCGACACATTCAAATNCCAGATGCTGCCCTGGTCTCAGGAGAAGT 703

## RESULT 4

BF523018

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

BF523018 604 bp mRNA linear EST 11-DEC-2000

UI-R-C2p-rg-f-10-0-UI.r1 UI-R-C2p Rattus norvegicus cDNA clone

UI-R-C2p-rg-f-10-0-UI 5', mRNA sequence.

BF523018 GI:11631033

EST.

Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.



```

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

1 (bases 1 to 604)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
CDNA Library Preparation: M.B. Soares Lab Clone Distribution:
clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LLNL (info@image.llnl.gov). IMAGE ID= 1793119
Seq primer: M13 Forward.
Location/Qualifiers
1. 604
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C2p-Ig-f-10-0-UI"
/clone_lib="UI-R-C2p"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pT730-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-C2p
library is a subtracted library derived from the UI-R-C1
library, which is a subtracted library derived from the
UI-R-C0 library. The UI-R-C0 library consisted of a
mixture of individually tagged normalized libraries
constructed from rat placenta, adult lung, brain, liver,
kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
embryo. The tag is a string of 3-5 nucleotides present
between the Not I site and the oligo-dT track which allows
identification of the library of origin of a clone within
the mixture. The subtracted library (UI-R-C2p) was
constructed as follows: PCR amplified cDNA inserts from
UI-R-C1 clones from which 3' ESTs had been derived was
used as a driver in a hybridization with the UI-R-C1
library in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library) was
purified by hydroxyapatite column chromatography
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
UI-R-C2p library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research 6:
791-806, 1996)"
BASE COUNT 166 a 132 c 167 g 139 t
ORIGIN
Query Match 12.6%; Score 440; DB 12; Length 604;
Best Local Similarity 85.4%; Pred. No. 2.2e-64;
Matches 504; Conservative 0; Mismatches 80; Indels 6; Gaps 1;

QY 2016 TGAAGAGAGGAGGACTCATGTCAGAACTCAGATGATGACCCAGCTGGGAAGCCACAGAA 2075
DB 9 TGAAGAGAGGCGCTCATGCTGAGCTCAAAATGATGACCCAGCTGGGACCATGACAA 68
QY 2076 TATTGTGAACCTGCTGGGGCGTGCACACTGTCAGGACCAATTTACTTGATTTTGAATA 2135
DB 69 CATCGTGAACCTGCTGGGGCGATGCACACTGTCAGGCGCCAGTGTACTTGATTTTGAATA 128
QY 2136 CTGTGCTGATGTCATCTTCTCAACTATCTTAAGAAGTAAAGAGAGAAAAATTTACAGAGC 2195
DB 129 TTGTTGCCATGGTGACCTGCTCAACTACCTTAAGACAAAGAGAGAAAGTTTCACAGGAC 188
QY 2196 TTGGACAGAGATTTTCAAGSAACACAAATTTTACGTTTTTACCACACTTTTCAATCACATCC 2255
DB 189 GTGACAGAGATTTTAAAGGAACATAATTTTACGTTTTTATACCCACCTTCCAGTCACATTC 248

FEATURES
Source
LOCUS AV713950 460 bp mRNA linear EST 11-OCT-2000
DEFINITION AV713950 DCB Homo sapiens cDNA clone DCBADE05 5', mRNA sequence.
ACCESSION AV713950
VERSION AV713950.1 GI:10795467
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 460)
AUTHORS Xu,X., Gu,J., Liu,F., Qu,J., Zhao,M., Li,Y., Huang,Q., Zhou,J.,
Song,H., Gu,Y., Yang,Y., Gao,G., Xiao,H., Li,N., Qian,B., Gao,X.,
Cheng,Z., Gu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu
G., Cheng,Z. and Han,Z.
Homo sapiens cDNA DCB clones
Unpublished (2000)
Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
source
1. 460
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DCBADE05"
/clone_lib="DCB"
/cell_type="dendritic cells"
/dev_stage="mature"
/lab_host="BM25.8"
/notes="Vector: pTriplex2; Site_1: sf1A; Site_2: sf1B"
BASE COUNT 152 a 95 c 96 g 117 t
ORIGIN
Query Match 12.5%; Score 437.8; DB 10; Length 460;
Best Local Similarity 99.5%; Pred. No. 5.8e-64;
Matches 439; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1930 ATGAACCCACAGCTTATGAATTAGCAAAACAGAGAGTCTCAATCCAGGTTGCCGTCAA 1989
DB 1 ATGAACCCACAGCTTATGAATTAGCAAAACAGAGAGTCTCAATCCAGGTTGCCGTCAA 60
QY 1990 ATGCTGAAGAAAAAGCAGACAGCTCTGAAAGAGAGGAGCACTCATGTCTCAGAACTCAAGATG 2049

```

```

|||||
61 ATGCTGAAAGAAAGAGAGAGAGCTCTCAAGAGAGGAGGACCTCATGTCAGAACTCAAGATG 120
QY
2050 ATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGCGTGACACTGTCA 2109
Db
121 ATGACCCAGCTGGGAAGCCAGAGAAATTTGTGAACCTGCTGGGGGCGTGACACTGTCA 180
QY
2110 GGACCAATTTACTGTGATTTTGAATCTGTTGCTATGCTGATCTCTCAACTATCTAAGA 2169
Db
181 GGACCAATTTACTGTGATTTTGAATCTGTTGCTATGCTGATCTCTCAACTATCTAAGA 240
QY
2170 AGTAAAGAGAGAAATTTACAGAGACTTGGACAGAGATTTTCAAGGAACAAATTTCACT 2229
Db
241 AGTAAAGAGAGAAATTTACAGAGACTTGGACAGAGATTTTCAAGGAACAAATTTCACT 300
QY
2230 TTTTACCCCACTTCCCAATCACATCCAAATTTCCAGCATGCTGCTTCAAGGAGAGTTTCA 2289
Db
301 TTTTACCCCACTTCCCAATCACATCCAAATTTCCAGCATGCTGCTTCAAGGAGAGTTTCA 360
QY
2290 ATACACCCGAGCTGGGATCAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGAT 2349
Db
361 ATACACCCGAGCTGGGATCAATCTCAGGGCTTCATGGGAATTCATTTCACTCTGAAGAT 420
QY
2350 GAAATGGAATGAAACCAA 2370
Db
421 GAAATGGAATGAAACCAA 441

```

## RESULT 6

```

AA120050
LOCUS
DEFINITION
mp88b02.r1 Soares_thymus.2NbMT Mus musculus cDNA clone IMAGE:576267
5', similar to gb:M64689 Mouse flk-2 mRNA, complete cds (MOUSE);,
mRNA sequence.

```

## ACCESSION

```

VERSION
AA120050.1 GI:1677559

```

## KEYWORDS

```

SOURCE
EST.

```

## ORGANISM

```

Mus musculus
house mouse.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognath; Muridae; Mus.
1 (bases 1 to 618)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Mei:350915

```

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

```

1..618
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:576267"
/clone_lib="Soares_thymus.2NbMT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

```

```

polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCCGGTTTTTTTTTTTTTTTTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 158 a 142 c 158 g 160 t
ORIGIN

```

```

Query Match 11.5%; Score 402.6; DB 9; Length 618;
Best Local Similarity 82.4%; Pred. No. 4.3e-58;
Matches 462; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 1582..GGGTCTCTGTCAGTGTGTCATACATAATTCCTTGGCAGATCTTGTGAGAGATTCCTT 1641
Db 1 GGGCTTCTGGTCAATGCTGCGTACAATTCATGGGCACGCTCTTGGCAAAACCATCTTT 60
QY 1642..TTAAACTCTCCAGGCCCTTCCCTTTCATCCAAAGACACATCTCATCTATGCAACAAT 1701
Db 61 TTAAACTCACCAGGCCCTTCCCTTTCATCCAAAGACACATCTCTTCTATGCGACCAT 120
QY 1702..GGTGTCTCTCTCTTCAATGTCCTTTTAAACCTGCTAATTTGTACAAAGTACAAAAAG 1761
Db 121 GGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
QY 1762..CAATTTAGTATGAAAGCCAGCTACAGATGTGTACAGTGACCGGTCTCTCTCAGATAATGAG 1821
Db 181 CAATTTAGTATGAAAGCCAGCTACAGATGTGTACAGTGACCGGTCTCTCTCAGATAATGAG 240
QY 1822..TACTTCTACGTTGATTTCAGAGATATGAATATGATCTCAATGGAGTTTCCAGAGAA 1881
Db 241 TACTTCTACGTTGATTTCAGAGATATGAATATGATCTCAATGGAGTTTCCAGAGAGAG 300
QY 1882..AATTAGAGTTTGGGAAGGTACTAGATCAGGTGCTTTTGGAAAAGTGTATGAACCAACA 1941
Db 301 AACTTAGAGTTTGGGAAGGTCTCGGGTCTCGGGTCTTTCGGGAGGGTGTATGAACGCCAG 360
QY 1942..GCTTATGAANTTAGCAAAACAGAGAGTCTCAATCCAGGTTCGCGTCAAAATGCTGAAGAA 2001
Db 361 GCTTATGCAATTAGTAAAACGGAGTCTCAATTCAGGTGGCGTGAAGATGCTAAGAGAG 420
QY 2002..AAAGCAGACACTCTGAAGAGAGAGCCTCATGTGAGAACTCAAGATGATGACCCAGCTG 2061
Db 421 AAAGCTGACAGCTGTGAAAAAGAGCTCTCATGTGCGAGCTCACAATGATGATGACCCAGCT 480
QY 2062..GGAAGCCAGAGATATTTGAACTGCTGAGCGGTGCGAGCTGCGAGCTGCGAGCTGCGAG 2121
Db 481 GGACACCATGACACATCTGTAATCTGCTGGGGCATGCACACTGTCAGGACCAATTTAC 540
QY 2122..TTGATTTTGAATGCTGTGC 2142
Db 541 CTGATTTTGAATGCTGTGC 561

```

## RESULT 7

## LOCUS

## DEFINITION

```

BM147250
TCAAP1010678 Pediatric acute myelogenous leukemia cell (FAB M1)
Baylor-HGSC project-TCOA Homo sapiens cDNA clone TCAAP1067, mRNA
sequence.

```

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

```

BM147250
BM147250.1 GI:17166911
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 455)
Wei, Y., Tsang, Y. T. M., Mei, G., Ku, J. M., Ali-Osman, F. R., Jr.,
Gunaratne, P. H., Muzny, D., Bouck, J., Gibbs, R. A. and Margolin, J. F.

```





QY	2515	ACCCAGGGAAAGTGGTGAAGATATGTGACTCTTTGGATTGGCTCCAGATATCATCATGAGTAT	2574
Db	2445	ACCAGCGGACATGTGGCCAAAGATTGGGAGACTTTGGAGCTGCTAGGACATCATGATGAC	2504
QY	2575	TCCAACATATGTTGTCAAGGGCAATGGCCGCTCTCCCTGTATAAATGGATGGCCCCGAAAGC	2634
Db	2505	TCCAACATATGTTCTCAAGGGCAATGGCCGCTCTCCCTGTATAAATGGATGGCCCCGAGAGC	2584
QY	2635	CTGTTTGAAGGCATCTACACCATTAAGATGATGCTGGTCATATGGAATATTAATCTCTGG	2694
Db	2565	ATCTTTGACTGCGCTCTACACAGATTCAGAGTGATGTGTGGTCTACGGCATCTCTCTCTGG	2624
QY	2695	GAATCTCTCACTGTGTGTGAATCCTTACCTTGGCATTCGCGTTGATGCTAACTCTCTAC	2754
Db	2625	GAGATCTTCTGCGTCTGTGCTGAACCCCTACCCGGCATCTAGTGACACAACTTCTCTAC	2684
QY	2755	AAACTGATTCAAATGATTTAAATGGATCAGCCATTTTATGCTACAGAAGAAATATAC	2814
Db	2685	AAACTGCTGAAGATGATACCAAAATGGCCAGCGCTATTGTGACCAAGAACAATATAC	2744
QY	2815	ATTATATGCAATCCTCTGGCTTTTGACTCAAGGAACGGCCATCCTTCC	2866
Db	2745	AGCATCATGCAGTCTCTGCTGGAGCTTGAGCCCTACCAGAAGACCCACTTCC	2796
RESULT 10			
BE511447/c			
LOCUS			
DEFINITION			
UI-M-BH3-atk-d-06-0-UI.r1 NIH_BMAP_M_S4 Mus musculus cDNA clone			
UI-M-BH3-atk-d-06-0-UI 5', mRNA sequence.			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
house mouse			
Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE			
AUTHORS			
TITLE			
Normalization and subtraction: two approaches to facilitate gene			
discovery			
Genome Res. 6 (9), 791-806 (1996)			
JOURNAL			
MEDLINE			
COMMENT			
Contact: Chin, H			
National Institute of Mental Health			
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD			
20892-9643, USA			
Tel: 301 443 1706			
Fax: 301 443 9890			
Email: mestr@mail.nih.gov			
cDNA library Preparation: M.B. Soares Lab Clone distribution:			
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It			
should be noted that Bento Soares is generating a small number of			
additional specialized non-redundant arrays of BMAP cDNAs whose			
availability will be considered under appropriate and limited			
collaborative arrangements			
Seq primer: M13 Reverse.			
FEATURES			
source			
1. .445			
/organism="Mus musculus"			
/strain="C57BL/6J"			
/db_xref="taxon:10090"			
/clone="UI-M-BH3-atk-d-06-0-UI"			
/clone.lib="NIH_BMAP_M_S4"			
/dev_stage="27-32 days"			
/lab_host="DH10B (Life Technologies)"			
/note="Vector: pT730-Pac (Pharmacia) with a modified			
polylinker; Site.1: Not I; Site.2: Eco RI; The			
NIH_BMAP_M_S4 library is a subtracted library of a series			
ultimately derived from a mixture of individually tagged			
normalized libraries from ten regions of the mouse brain			
(cerebellum, brain stems, olfactory bulbs, hypothalamus,			

cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH\_BMAP\_M\_S4, NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, NIH\_BMAP\_M\_S3.1, NIH\_BMAP\_M\_S2, NIH\_BMAP\_M\_S1. The subtracted library (NIH\_BMAP\_M\_S4) was constructed as follows: PCR-amplified cDNA inserts from NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, and NIH\_BMAP\_M\_S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH\_BMAP\_M\_S3.3, NIH\_BMAP\_M\_S3.2, and NIH\_BMAP\_M\_S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH\_BMAP\_M\_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)."

BASE COUNT 96 a 123 c 111 g 115 t  
ORIGIN

Query Match 9.3%; Score 325.6; DB 10; Length 445;  
Best Local Similarity 83.3%; Pred. No. 4.3e-45;  
Matches 370; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 637 GTGGATGGTGTTCGCGATTCACAGGGGGAAGCTGTAAAGAAAGTCCAGCTGTT 696  
DB 444 GTGGAGTGGTGTCTGTCAGCTCCACAGGGAAGCTGTAAAGAAAGGCGCTGCTGTT 385  
QY 697 GTTAAAGAGGAGGAAAAAGTCTTCATGAATTTTGGGACGACATAGGCTGTGCC 756  
DB 384 GTCAGAAAGAGGAAAGGTACTTCATGAGTTTTCGGAACACATCATGCTGTGCT 325  
QY 757 AGAATGAATGCGGAGGAATGACACAGGCTGTTCACAAATAGATCTAAATCAAACTCCT 816  
DB 324 AGAATGCACTGGCGCGAATSCACCAAGCTGTTCACCATAGATCTAAACAGGCTCCT 265  
QY 817 CAGACACATGTCACCAATATTCTTAAAGTAGGGGAACCCCTTATGGATAAGGTGCAA 876  
DB 264 CAGACACACTGCGCCAGTATTCTTGAAGTAGGGGAACCCCTTATGGATCAGGTGTAAG 205  
QY 877 GCTGTTCTATGTAACCATGATTCGGGCTCACCTGGGAATTAGAAACAAAGCAGCTCGAG 936  
DB 204 GCATCCATGTAACCATGATTCGGGCTCACCTGGGAGCTGGAACAAAGCCCTGGAG 145  
QY 937 GAGGCAACTACTTTGAGATGAGTACCTATTCAACAAACAGAACTATGATACGATTCTG 996  
DB 144 GAGGCGAGCTACTTTGAGATGAGTACCTACTCCAAACAGGACCATGATTCGATCTC 85  
QY 997 TTTGCTTTTGTATCATCAGTGGCAAGAACGACACCGGATACATCTGTTCTCTCTCA 1056  
DB 84 TTGGCTTTGTGTCTTCCGGAAGAACGACACCGGATATACACCTGCTCTCTCTCA 25  
QY 1057 AAGCATCCAGTCAATCAGCTTG 1080  
DB 24 AAGCACCACCGCAGTCAAGCTTG 1

## RESULT 11

BC026713  
LOCUS 5174 bp mRNA linear HTC 07-AUG-2002  
DEFINITION Mus musculus, clone IMAGE:5008623, mRNA.  
ACCESSION BC026713  
VERSION BC026713.1 GI:19934327  
KEYWORDS HTC.  
SOURCE house mouse.  
-ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 5174)  
AUTHORS Strausberg, R.

TITLE  
JOURNALREMARK  
COMMENT

Direct Submission  
Submitted (04-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaphs-remail.nih.gov](mailto:cgaphs-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
Gunnaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 45 Row: b Column: 17  
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis  
This clone has the following problem: no 5' EST match.

FEATURES  
source

Location/Qualifiers  
1. 5174  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/map="FVB/N-3"  
/clone="IMAGE:5008623"  
/tissue\_type="Mammary tumor. MMTV-LTR/INT3 model. 5 month old mouse. Taken by biopsy."  
/clone\_lib="NCI\_CGAP\_Mam2"  
/lab\_host="DH10B"  
/note="vector: pCMV-SPORT6"  
BASE COUNT 1356 a 1232 c 1246 g 1340 t  
ORIGIN

Query Match 9.1%; Score 318.6; DB 11; Length 5174;  
Best Local Similarity 57.1%; Pred. No. 2.3e-44;  
Matches 601; Conservative 0; Mismatches 449; Indels 3; Gaps 1;  
QY 1816 AATGAGTACTCTACGTTGATTTTCAGAGATATGAATATGATCTCAATGGAGTTTCCA 1875  
DB 1749 AACAAATTATGTTTACATAGACCCGACGCAACTTCTTATGATCACAATGGAGTTTCCC 1808  
QY 1876 AGAGAAATTTAGAGTTTGGGAAGGTACTAGGATCAGGTGCTTTTGGAAAAGTGAAGAC 1935  
DB 1809 AGAAACAGGCTGAGTTTGGGAAGACATTTGGAGCTGGTCTTCGGGAAGGTCGTTGAG 1868  
QY 1936 GCAACAGCTTATGGAATTAGCAAAACAGAGTCTCAATCCAGGTTCCGTCACAAATGCTG 1995  
DB 1869 GCCACTCATATGCTGTTGATTAAGTCGGATGCTGCCATGACAGTTGCCGTGAAGATGCTC 1928  
QY 1996 AAAGAAAAGACAGACAGCTCTGAAAGAGAGGCACTCATGTGAGAACTCAAGATGATGACC 2055  
DB 1929 AAACCAAGTCCCATTTTAAACAGAAAGAGAGGCCCTATGTCGGAACCTGAAGTCTGAGC 1988  
QY 2056 CAGCTGGGAAGCCACAGAGATATTGTGAACCTGCTGGGGCGTGCACACTGTCTAGGACCA 2115  
DB 1989 TACCTGGGCAATCACATGAATATTGTGAACCTGCTGGCGCATGCGGTGGAGGGGCC 2048  
QY 2116 ATTTACTTGAATTTTGAATACTGTTGCTATGGTATCTTCTCACTATCTAAGAAGTAAA 2175  
DB 2049 ACCCTGCTCATACAGATATTGTTGCTATGGTATCTTTTGAATTTTGTGAGAAGGAAG 2108  
QY 2176 AGAGAAAAATTTTACAGGAGCTTGGACAGAGATTTTCAAGGAACACAAATTTTCAGTTTATC 2235  
DB 2109 CGTGACTCGTTTATTTTCTCAAGCAAGAGAGCAGGACGACGAGGCGGCACITTTATAAGAAC 2168





Db 127 CAACCAAGACCTCCCTGTGTGATCAAGTGTGTTTAAATCAGTCATGAGAACAAATGGCTCAT 186

Qy 194 CAGTGGGAAGTCATCATATATATCCATGATCCAGTATCCCGGAAGACCTCGGGTGTG 253

Db 187 CAGCGGAAGACCATCATCTGATCCAGTATGTCGAGGATCCCGAAGACCTCCAGTGTG 246

Qy 254 CGTTGAGACCCAGAGCTCAGGACAGTGTACGAGCTGCGCTGTGGAAGTGTAT 313

Db 247 CCGCGAGCGCCAGAGTGAAGGACGGTATATGAAGCGGCCACCGTGGAGTGGCCGAGT 306

Qy 314 CTGCTTCATCACACTGCAAGTGTGTCGATGCCCGAGGAAACATTTCTTCTCTCTGGG 373

Db 307 CTGGTCCATCACCTCCAGTGTGCTGCGCCACCCAGGGGACCTTTCTTCTCTGGG 366

Qy 374 TCTTTAAGCAGCTCCCTGAAATTCAGGCGACACATTTTGAATTTACAAACAGAGAGTGT 433

Db 367 TCTTTAAGCAGCTCCCTGGCTGCCAGCGCACTTTGATTTACAAACAGAGAGATCG 426

Qy 434 TTTCCATGCTCATTTCGAATATGACAGAACCCCAAGCTGGAGATACCTATTTATTC 493

Db 427 TTTCCATGGCCATCTTGAAGCTGTACAGAGACCCAGGAGGAGATACCTACTCCATATTC 486

Qy 494 AGAGTGAAGCTACCAATTTACACAAATTTGTTTACAG 529

Db 487 AGAGCGAAGCGCC-ACTACACAGTACTGTTTACAG 521

## RESULT 13

BM144393/c

## LOCUS

## DEFINITION

TCAPID10678 Pediatric acute myelogenous leukemia cell (FAB M1)  
Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAAP1067, mRNA

## ACCESSION

BM144393

## VERSION

EST.

## KEYWORDS

human.

## SOURCE

Homo sapiens.

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 326)

Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, F.R. Jr.,

Gunaratne, P.H., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.

Pediatric Leukemia cDNA Sequencing Project (2001)

Unpublished (2001)

Contact: Dr. Judith F. Margolin

Texas Children's Cancer Center and Human Genome Sequencing Center

at Baylor College of Medicine

1102 Bates, MC3-3320 Houston, TX 77030, USA

Tel: 832-824-4536

Fax: 832-825-4038

Email: clones@txccc.org

Seq primer: M13 primer.

Location/Qualifiers

1..326

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="TCAAP1067"

/clone\_lib="Pediatric acute myelogenous leukemia cell (FAB

M1) Baylor-HGSC project-TCAA"

/sex="male"

/tissue\_type="leukopheresis"

/cell\_type="myeloid cell"

/dev\_stage="pediatric 6 years"

/lab\_host="DH10B"

/notes="Vector: lambda pSB; Site 1: BamHI; Site 2: EcoRI;

First strand cDNA was primed with an anchored

XhoI-oligo(dT) primer [5'GGAGACTCGAGCGCGCGAGGAG(T)VN

3'; V-A, C, G, N-A, C, G, T] and then dG tailed. Second strand

was primed with a BamHI-dC primer

[5'AGAGAGCTCGATCCGCGCGCAATAATAAT(C) 3'].

Double-stranded cDNA was then digested with BamHI and XhoI

and directionally cloned into the BamHI and SalI sites of

## FEATURES

source

lambda pSB vector. Library went through one round of normalization. Library was constructed by Wei Yu at RIKEN of Japan (Carninci P. Westover A. Nishiyama Y. Ohsumi T, Itoh M, Nagaoka S. Sasaki, Okazaki Y, Muramatsu M, Schneider C, Hayashizaki Y, High efficiency selection of full-length cDNA by improved biotinylated cap trapper., DNA Res 4: 1, 61-6, Feb 28, 1997)"

BASE COUNT 111 a 56 c 45 g 114 t  
ORIGIN

Query Match 8.7%; Score 305.4; DB 13; Length 326;  
Best Local Similarity 97.9%; Pred. No. 1.2e-41;  
Matches 320; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

Qy 3163 TTTCTCTAGAACCGCTCTGCTTTACTCTCTTTTCAAGGACACTTTCTAAATCAAT 3222

Db 326 TTTCTCTACAGCTGTCTGCGTTTACTCTTTTACAAAGA-ACTTTGTAAATCAAT 268

Qy 3223 CATCTGTCAACAGCAGGAGGAGCTGATATGAACCTTTATGGAGCATTGATCTGCATC 3282

Db 267 CATCTGTCAACAGCAGGAGGAGCTGATATGAACCTTTATGGAGCATTGATCTGCATC 208

Qy 3283 CAAGCCCTTCAGGCCGCTTGAGTGAATTCGTGACCTGAAGTACAGTATATCTTGA 3342

Db 207 CAAGCCCTTCAGGCCGCTTGAGTGAATTCGTGACCTGAAGTACAGTATATCTTGA 148

Qy 3343 AATACATAAACAACAGCATTTCCTAAGGAGAGCTAATATGATTTTAACTATGT 3402

Db 147 AATACATAAACAACAGCATTTCCTAAGGAGAGCTAATATGATTTTAACTATGT 88

Qy 3403 TTTAAATAATATGTAATTTTTCAGCTATTTAGTGATATATTTATGGTGGGAATAA 3462

Db 87 TTTAAATAATATGTAATTTTTCAGCTATTTAGTGATATATTTATGGTGGGAATAA 28

Qy 3463 ATTTCTACTACAGAAAAA 3489

Db 27 ATTTCTACTACAGAAAAA 1

## RESULT 14

AW772610/c

## LOCUS

hn77b03.x1

NCI-CGAP Kid11 Homo sapiens

similar to gb:U02687 FL CYTOKINE RECEPTOR PRECURSOR (HUMAN);, mRNA

sequence.

AW772610

EST.

AW772610.1

GI:7704608

human.

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 646)

Email: cgapbs-r@mail.nih.gov

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

DNA sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL, send email to:

info@image.llnl.gov

Seq primer: -400P from Gibco

High quality sequence stop: 462.

Location/Qualifiers

1..646

/organism="Homo sapiens"

/db\_xref="taxon:9606"

Insert Length: 1584 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 431.

FEATURES  
Source  
Location/Qualifiers  
1..659  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:2542486"  
/clone\_lib="NCI\_CGAP\_Gas4"  
/tissue\_type="poorly differentiated adenocarcinoma with  
signet ring cell features"  
/lab\_host="DH10B"  
/note="Organ: stomach; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.69 kb. Life Technologies catalog #:  
11549-011"

BASE COUNT 204 a 131 c 119 g 205 t  
ORIGIN

Query Match 8.1%; Score 283.4; DB 10; Length 659;  
Best Local Similarity 98.0%; Pred. No. 4.5e-38;  
Matches 287; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3185 TTACTCTTGTTCCTCAAGGACCTTTTGTAAATCAATCATCTGTGCACAGGAGG 3244  
Db 639 TTACTCTTGTTCCTCAAGGACCTTTTGTAAATCAATCATCTGTGCACAGGAGG 580  
QY 3245 AGCTGATAATGAACCTTTATTGGAGCATTTGATCTGCATCCAGGCTTCTCAGGCGGCTT 3304  
Db 579 AGCTGATAATGAACCTTTATTGGAGCATTTGATCTGCATCCAGGCTTCTCAGGCGGCTT 520  
QY 3305 GAGTGAATCTGTACCTGAACTACAGTATATCTTGTAAATACATAAAACAAAGCATTT 3364  
Db 519 GAGTGAATCTGTACCTGAACTACAGTATATCTTGTAAATACATAAAACAAAGCATTT 460  
QY 3365 TGCTAAGGAGAACTAATATGATTTTAAAGTCTATGTTTAAATATATATGTAATTTT 3424  
Db 459 TGCTAAGGAGAACTAATATGATTTTAAAGTCTATGTTTAAATATATATGTAATTTT 400  
QY 3425 TCAGCTATTTAGTCATATATTTTATGGTGGGAATAAAATTTCTACTACAGAA 3477  
Db 399 TCAGCTATTTAGTCATATATTTTATGGTGGGAATAAAATTTCTACTACAGAA 347

Search completed: May 25, 2003, 11:51:19  
Job time : 4599.52 secs

/clone="IMAGE:3033869"  
/clone\_lib="NCI\_CGAP\_Kid11"  
/lab\_host="DH10B"  
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with  
a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;  
plasmid DNA from the normalized library NCI\_CGAP\_Kid3 was  
prepared, and ss circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(cloneIDs 1322376-1323911, 1456007-1456775, and  
1500552-1502855). Subtraction by Bento Soares and M.  
Fatima Bonaldo."

BASE COUNT 201 a 130 c 116 g 199 t  
ORIGIN

Query Match 8.6%; Score 299.6; DB 10; Length 646;  
Best Local Similarity 98.7%; Pred. No. 8.7e-41;  
Matches 302; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3172 AAGCGCTGCGTTTACTCTTGTTCCTCAAGGACCTTTTGTAAATCAATCATCTCTGC 3231  
Db 646 AAGCTGCTGCGTTTACTCTTGTTCCTCAAGGACCTTTTGTAAATCAATCATCTCTGC 597  
QY 3232 ACAAGCAGGAGGAGCTGATATGAACCTTTATTGGAGCATTTGATCTGCATCCAGGCGCTT 3291  
Db 586 ACAAGCAGGAGGAGCTGATATGAACCTTTATTGGAGCATTTGATCTGCATCCAGGCGCTT 527  
QY 3292 CTCAGCGCGCTTGAAGTGAATGTGTACCTGAGTACAGTATATCTTGTAAATACATAA 3351  
Db 526 CTCAGCGCTGCTGAGTGAATGTGTACCTGAGTACAGTATATCTTGTAAATACATAA 467  
QY 3352 AACAAAGCATTTTGTCTAAGGAGAACTAATATGATTTTAAAGTCTATGTTTAAATA 3411  
Db 466 AACAAAGCATTTTGTCTAAGGAGAACTAATATGATTTTAAAGTCTATGTTTAAATA 407  
QY 3412 ATATGTAATTTTTCAGCTATTTAGTATATATTTTATGGTGGGAATAAAATTTCTACT 3471  
Db 406 ATATGTAATTTTTCAGCTATTTAGTATATATTTTATGGTGGGAATAAAATTTCTACT 347  
QY 3472 ACAGAA 3477  
Db 346 ACAGAA 341

RESULT 15  
AW057705/c  
LOCUS  
DEFINITION  
wx02c12.x1 NCI\_CGAP\_Gas4 Homo sapiens cDNA clone IMAGE:2542486 3,  
similar to gb:U02687 FL CYTOKINE RECEPTOR PRECURSOR (HUMAN);, mRNA  
sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORIGIN  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 659)  
REFERENCE  
AUTHORS  
TITLE  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
CONTACT: Robert Strausberg, Ph.D.  
Email: cgaps-r@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:18:14 ; Search time 61.0307 seconds  
(without alignments)  
2168.053 Million cell updates/sec

Title: US-09-919-408-4  
Perfect score: 5274  
Sequence: 1 MPALARDAGTVPLLVFSAM .....PFSREMDLGLSPQAQVEDS 993

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

1: /SID52/gcgdata/genseq/genseq-emb1/AA1980.DAT:\*

2: /SID52/gcgdata/genseq/genseq-emb1/AA1981.DAT:\*

3: /SID52/gcgdata/genseq/genseq-emb1/AA1982.DAT:\*

4: /SID52/gcgdata/genseq/genseq-emb1/AA1983.DAT:\*

5: /SID52/gcgdata/genseq/genseq-emb1/AA1984.DAT:\*

6: /SID52/gcgdata/genseq/genseq-emb1/AA1985.DAT:\*

7: /SID52/gcgdata/genseq/genseq-emb1/AA1986.DAT:\*

8: /SID52/gcgdata/genseq/genseq-emb1/AA1987.DAT:\*

9: /SID52/gcgdata/genseq/genseq-emb1/AA1988.DAT:\*

10: /SID52/gcgdata/genseq/genseq-emb1/AA1989.DAT:\*

11: /SID52/gcgdata/genseq/genseq-emb1/AA1990.DAT:\*

12: /SID52/gcgdata/genseq/genseq-emb1/AA1991.DAT:\*

13: /SID52/gcgdata/genseq/genseq-emb1/AA1992.DAT:\*

14: /SID52/gcgdata/genseq/genseq-emb1/AA1993.DAT:\*

15: /SID52/gcgdata/genseq/genseq-emb1/AA1994.DAT:\*

16: /SID52/gcgdata/genseq/genseq-emb1/AA1995.DAT:\*

17: /SID52/gcgdata/genseq/genseq-emb1/AA1996.DAT:\*

18: /SID52/gcgdata/genseq/genseq-emb1/AA1997.DAT:\*

19: /SID52/gcgdata/genseq/genseq-emb1/AA1998.DAT:\*

20: /SID52/gcgdata/genseq/genseq-emb1/AA1999.DAT:\*

21: /SID52/gcgdata/genseq/genseq-emb1/AA2000.DAT:\*

22: /SID52/gcgdata/genseq/genseq-emb1/AA2001.DAT:\*

23: /SID52/gcgdata/genseq/genseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5274	100.0	993	16 AAR67816	Flk2 receptor prot
2	5274	100.0	993	16 AAR67536	Human flk-2. Homo
3	5274	100.0	993	17 AAR97419	Murine foetal live
4	5274	100.0	993	18 AAW19873	Human flk-2 recept
5	5274	100.0	993	20 AAY08617	Human flk-2 protei
6	5271	99.9	993	14 AAR37503	Human flk-2. Homo
7	5266	99.8	993	14 AAR44995	Human flk-2 recept
8	5266	99.8	993	16 AAR75961	Human STR-1. Homo
9	5265	99.8	993	16 AAR81869	Human Flk2/flt3 ty
10	5262	99.8	1167	14 AAR31376	Human flk-2. Homo

11	5159	97.8	983	19 AAW63588	Human receptor typ
12	5157.5	97.8	986	19 AAW63587	Human receptor typ
13	5157.5	97.8	986	19 AAW63589	Human receptor typ
14	5153.5	97.7	994	19 AAW63586	Human receptor typ
15	4533.5	86.0	1000	16 AAR81868	Flk2/flt3 tyrosine
16	4429.5	84.0	992	14 AAR37502	Murine flk-2. Mus
17	4429.5	84.0	992	16 AAR67815	Flk2 receptor prot
18	4429.5	84.0	992	16 AAR67535	Murine flk-2. Mus
19	4429.5	84.0	992	17 AAR97418	Human foetal liver
20	4429.5	84.0	992	18 AAW19874	Murine flk-2 recep
21	4429.5	84.0	992	20 AAY08616	Murine flk-2 prote
22	4421.5	83.8	992	14 AAR44994	Murine flk-2 recep
23	4421.5	83.8	992	14 AAR31375	Murine flk-2. Mus
24	4404.5	83.5	992	13 AAR28038	Murine flk-2. Mus
25	3323	63.0	665	19 AAW63585	Human receptor typ
26	2747	52.1	749	19 AAW78002	Protein pMON32390.
27	1640	31.1	481	15 AAR47579	Soluble Flk-2. Mu
28	1286	24.4	977	21 AAY51322	Bovine c-Kit bk-1
29	1266	24.0	975	22 AAE07144	Murine Kit/stem ce
30	1266	24.0	975	22 AAE07148	Mutant murine Kit/
31	1251	23.7	976	22 AAE07145	Human Kit/stem cel
32	1251	23.7	976	22 AAE07149	Mutant human Kit/s
33	1251	23.7	976	22 AAE00375	Human stem cell gr
34	1219.5	23.1	1055	22 ABG15479	Novel human diagno
35	1218.5	23.1	972	23 AAU11936	Colony stimulating
36	1217.5	23.1	972	23 AAU11934	Colony stimulating
37	1217.5	23.1	972	23 AAU11940	Colony stimulating
38	1216.5	23.1	972	23 AAU79039	Human macrophage c
39	1216.5	23.1	972	23 AAU11941	Colony stimulating
40	1216.5	23.1	972	23 AAU73585	Colony stimulating
41	1215.5	23.0	972	23 AAU11935	Colony stimulating
42	1215.5	23.0	972	23 AAU11937	Colony stimulating
43	1211.5	23.0	972	23 AAU11938	Colony stimulating
44	1207.5	22.9	972	23 AAU11939	Colony stimulating
45	1189.5	22.6	2129	22 ABG15478	Novel human diagno

ALIGNMENTS

RESULT 1	
AAR67816	
ID AAR67816 standard; Protein; 993 AA.	
XX AAR67816;	
AC AAR67816;	
XX AAR67816;	
DT 18-AUG-1995 (first entry)	
XX Flk2 receptor protein-tyrosine-kinase.	
DE Flk2 receptor protein-tyrosine-kinase.	
XX Human Flk2; receptor protein-tyrosine-kinase; primitive	
KW hematopoietic cell; fetal liver kinase; diagnostic ligand	
KW isolation; bone marrow disease therapy.	
XX Homo sapiens.	
OS Homo sapiens.	
XX Key	Location/Qualifiers
FT Peptide	1..27
FT Protein	/note= "signal peptide"
FT Protein	28..993
FT Protein	/note= "mature protein"
FT Domain	28..343
FT Domain	/note= "extracellular receptor domain"
FT Domain	544..563
FT Domain	/note= "transmembrane region"
FT Domain	564..993
FT Domain	/note= "intracellular catalytic domain"
XX WO9500554-A.	
PN WO9500554-A.	
XX Human flk-2 recept	
PD 05-JAN-1995.	
XX Human Flk2/flt3 ty	
PF 17-JUN-1994; 94WO-0506944.	

```
XX 18-JUN-1993; 93US-0080244.
PR 21-JUN-1993; 93US-0081508.
PR 23-NOV-1993; 93US-0157490.
XX
PA (UYPR-) UNIV PRINCETON.
XX
PI Lemischka IR;
XX
XX WPI; 1995-052014/07.
XX N-PSDB; AAQ81013.
XX
PT Ligand for receptor protein tyrosine kinase - useful for the
PT stimulation of primitive hematopoietic stem cells causing
PT proliferation and/or differentiation
XX
PS Disclosure; Fig 1b; 131pp; English.
XX
XX The sequence corresponds to a human flk2 (fetal liver kinase)
CC receptor protein-tyrosine-kinase, which is expressed in primitive
CC hematopoietic cells but not in mature hematopoietic cells. the
CC protein is useful in isolation of receptor ligands, which have
CC applications in diagnosis of bone marrow disorders and in
CC stimulating proliferation and/or differentiation of primitive
CC hematopoietic stem cells.
XX
SQ Sequence 993 AA;
Query Match 100.0%; Score 5274; DB 16; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPALARDAGTVPLLVFSAMIFGTITNQDLPIVKVLIHKNNDSSVGKSSSPYMVSESP 60
DB 1 MPALARDAGTVPLLVFSAMIFGTITNQDLPIVKVLIHKNNDSSVGKSSSPYMVSESP 60
QY 61 EDLGCALRPSQSGTVYAAAVEVDVSASITLQVLVDPAGNISCLWFKHSSLCNCQPHFDL 120
DB 61 EDLGCALRPSQSGTVYAAAVEVDVSASITLQVLVDPAGNISCLWFKHSSLCNCQPHFDL 120
QY 121 QNRGVVSMVILKMTETQAGEYLLFIOSEATNTYILFTVSIRNTLLYLRPRYFRKMENQD 180
DB 121 QNRGVVSMVILKMTETQAGEYLLFIOSEATNTYILFTVSIRNTLLYLRPRYFRKMENQD 180
QY 181 ALVCISESVPEIVWLCDSQGESCKEESPAVVVKEKVLHFLFGTDIRCCARNELGRE 240
DB 181 ALVCISESVPEIVWLCDSQGESCKEESPAVVVKEKVLHFLFGTDIRCCARNELGRE 240
QY 241 CTRFLTIDLNOTPQTTLPOLFLKVGPLWIRCAVHVNHGFGTLWELENKALEEGNYFEM 300
DB 241 CTRFLTIDLNOTPQTTLPOLFLKVGPLWIRCAVHVNHGFGTLWELENKALEEGNYFEM 300
QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSDYE 360
DB 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSDYE 360
QY 361 IDQYEEFCFSVRKAYPQIRCTWTFGRKSPCEQGLDNGYSISKFCNHRKHQCEYIFHA 420
DB 361 IDQYEEFCFSVRKAYPQIRCTWTFGRKSPCEQGLDNGYSISKFCNHRKHQCEYIFHA 420
QY 421 ENDDAQFTKMTILNIRKPPVLAESAQSACFSDGYPLPSWTWKKCSKSPNCTEITE 480
DB 421 ENDDAQFTKMTILNIRKPPVLAESAQSACFSDGYPLPSWTWKKCSKSPNCTEITE 480
QY 481 GWNRRKANRVFCQWSSSTLNHSEATKGLVKCCAYNSLGTSCETILLNSPGFPFIQD 540
DB 481 GWNRRKANRVFCQWSSSTLNHSEATKGLVKCCAYNSLGTSCETILLNSPGFPFIQD 540
QY 541 NISFYATIGVCLLFIIVVLTLLCHKYPKOPRYESQLQWQVGTSSDNEYFYVDFREYED 600
DB 541 NISFYATIGVCLLFIIVVLTLLCHKYPKOPRYESQLQWQVGTSSDNEYFYVDFREYED 600
QY 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIQVAVKMLKEKADSSEREALMS 660
```

```
DB 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIQVAVKMLKEKADSSEREALMS 660
QY 661 ELKMMTQLGSHENIVNLLGACTLSGGPIYLIFEYCCYGDLLNLYLSRKREKPHRTWTEIFE 720
DB 661 ELKMMTQLGSHENIVNLLGACTLSGGPIYLIFEYCCYGDLLNLYLSRKREKPHRTWTEIFE 720
QY 721 HNFSEFYPTQSHSPNSMPGSRVQIHPDSQOISGLHGNSEFHSDEIEYENOKRLEEEEDL 780
DB 721 HNFSEFYPTQSHSPNSMPGSRVQIHPDSQOISGLHGNSEFHSDEIEYENOKRLEEEEDL 780
QY 781 NVLTFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLTTHGKVVVKICDFGLARDINSDS 840
DB 781 NVLTFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLTTHGKVVVKICDFGLARDINSDS 840
QY 841 NYVVRGNARLPVKWMAPESLFEGIIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
DB 841 NYVVRGNARLPVKWMAPESLFEGIIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
QY 901 LIONGFKMDQPPYATEEIIYIMOSWAFDSKRPSPFNLTSLGCOLADAEAEAMYQNVDG 960
DB 901 LIONGFKMDQPPYATEEIIYIMOSWAFDSKRPSPFNLTSLGCOLADAEAEAMYQNVDG 960
QY 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993
DB 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993
RESULT 2
AAR67536
ID AAR67536 standard; Protein; 993 AA.
XX AAR67536;
XX 04-JUL-1995 (first entry)
XX Human flk-2.
XX DE
XX KW Fetal liver kinase-2; flk-2; protein tyrosine-kinase receptor;
XX KW hematopoiesis; stem cell.
XX OS Homo sapiens.
XX FH Key
XX FT Peptide 1..27 Location/Qualifiers
XX FT Domain 28..543 /label= Sig_peptide
XX FT Region 544..563 /label= Extracellular_receptor_domain
XX FT Domain 564..993 /label= Transmembrane_region
XX US5367057-A.
XX PN
XX PD 22-NOV-1994.
XX PF 02-APR-1991; 91US-0679666.
XX PR 02-APR-1991; 91US-0679666.
XX PR 28-JUN-1991; 91US-0728913.
XX PR 15-NOV-1991; 91US-0793065.
XX PR 24-DEC-1991; 91US-0813593.
XX PR 26-JUN-1992; 92US-0906397.
XX PR 12-NOV-1992; 92US-0975049.
XX PR 19-NOV-1992; 92US-0977451.
XX PR 30-APR-1993; 93US-0055269.
XX (UYPR-) UNIV PRINCETON.
XX PA
XX PI Lemischka IR;
XX PI WPI; 1995-005894/01.
XX DR
```



XX The present sequence is murine foetal liver kinase 2 (flk-2),  
 CC a protein tyrosine kinase. Isolated antibodies, pref. monoclonal,  
 CC raised against the extracellular portion of flk-2 can be used to  
 CC assay for flk receptors on the surface of primitive haematopoietic  
 CC stem cells, and to isolate positive cells. The antibodies can also  
 CC be used as, or to obtain ligands, which stimulate the proliferation  
 CC and/or differentiation of stem cells. The ligands can be used, e.g.  
 CC for treating anaemia, or bone marrow damage resulting from cancer  
 CC chemotherapy, or radiation.  
 XX  
 SQ Sequence 993 AA;

Query Match 100.0%; Score 5274; DB 17; Length 993;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPVLLVVFSAIFGTITNODLPVKVILNHKNNDSSVGKSSSYPMVSESP 60  
 DB 1 MPALARDAGTVPVLLVVFSAIFGTITNODLPVKVILNHKNNDSSVGKSSSYPMVSESP 60

QY 61 EDLGCALRPOSSGTVEAAVVDVSAITLQVLVDAPGNISCLWFKHSLNCQPHFDL 120  
 DB 61 EDLGCALRPOSSGTVEAAVVDVSAITLQVLVDAPGNISCLWFKHSLNCQPHFDL 120

QY 121 QNRGVVSMVTLKMTETQAGEYLLFIOSEATNYTLFTVSRNTLLYLRPRPKMENQD 180  
 DB 121 QNRGVVSMVTLKMTETQAGEYLLFIOSEATNYTLFTVSRNTLLYLRPRPKMENQD 180

QY 181 ALVCISESPEPTVEWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
 DB 181 ALVCISESPEPTVEWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240

QY 241 CTRLFTIDLNOTPOTLPQLFLKVGPELWIRKAVHNVHGFGLTWELNKALEEGNYFEM 300  
 DB 241 CTRLFTIDLNOTPOTLPQLFLKVGPELWIRKAVHNVHGFGLTWELNKALEEGNYFEM 300

QY 301 STYSTNRTMIRILFAFVSSVARNDGYTCCSSKHPSQSALVTIVGKGFNATNSSEDI 360  
 DB 301 STYSTNRTMIRILFAFVSSVARNDGYTCCSSKHPSQSALVTIVGKGFNATNSSEDI 360

QY 361 IDQYEFCEFSVREKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGGEYIFHA 420  
 DB 361 IDQYEFCEFSVREKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGGEYIFHA 420

QY 421 ENDDAFTKMTLNIRKPOVLAESAASQSCFSDGYPLPSWTWKKCDKSPNCTEITE 480  
 DB 421 ENDDAFTKMTLNIRKPOVLAESAASQSCFSDGYPLPSWTWKKCDKSPNCTEITE 480

QY 481 GVNRRKANRVFGQWSSSTLNSEATKGLVKCCAVNSLGTSCETILLNSPGPFPIQD 540  
 DB 481 GVNRRKANRVFGQWSSSTLNSEATKGLVKCCAVNSLGTSCETILLNSPGPFPIQD 540

QY 541 NISFYATIGVCLLFIVVLTLLCHKVKQRYESQLOMVOVQVSSDNEYFYVDFREYEYD 600  
 DB 541 NISFYATIGVCLLFIVVLTLLCHKVKQRYESQLOMVOVQVSSDNEYFYVDFREYEYD 600

QY 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIOVAVKMLKEKADSSREALMS 660  
 DB 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIOVAVKMLKEKADSSREALMS 660

QY 661 ELKMMTOLGSHENIVNLGACTLSGPIYLIIFYCCYGDLLNLYRSKREKPHRTWTEIFKE 720  
 DB 661 ELKMMTOLGSHENIVNLGACTLSGPIYLIIFYCCYGDLLNLYRSKREKPHRTWTEIFKE 720

QY 721 HNF5FTPTFOSHNSMPPGSRVQIHPDSQISGLHGNFSHSEDEIEYENOKRLEEBEDL 780  
 DB 721 HNF5FTPTFOSHNSMPPGSRVQIHPDSQISGLHGNFSHSEDEIEYENOKRLEEBEDL 780

QY 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVYKICDFGLARDINSDS 840  
 DB 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVYKICDFGLARDINSDS 840

QY 841 NYVVRGNARLPVKWMAPESLFEGYITTKSDVWSYGILLWEIFSLGVNYPGIPVDANEYK 900  
 DB 841 NYVVRGNARLPVKWMAPESLFEGYITTKSDVWSYGILLWEIFSLGVNYPGIPVDANEYK 900

QY 901 LIQNGFKMDQPFYATEEYIIIMQSCWAFDSKRSPFNLTSLFLGCOLADAEAMYNQVNDG 960  
 DB 901 LIQNGFKMDQPFYATEEYIIIMQSCWAFDSKRSPFNLTSLFLGCOLADAEAMYNQVNDG 960

QY 961 RVSECPHTYQNRPRFSREMDLGLLSPQAQVEDS 993  
 DB 961 RVSECPHTYQNRPRFSREMDLGLLSPQAQVEDS 993

RESULT 4  
 AAW19873  
 ID AAW19873 standard; Protein; 993 AA.  
 XX  
 AC AAW19873;  
 XX  
 DT 19-AUG-1997 (first entry)  
 XX  
 DE Human flk-2 receptor.  
 XX  
 KW Human; fetal liver kinase 2; flk2; receptor protein tyrosine kinase;  
 KW ptk; liver; spleen; thymus; adult; brain; bone marrow; primitive portion;  
 KW haematopoietic hierarchy; extracellular domain; soluble form; ligand;  
 KW proliferation; differentiation; mammalian; haematopoietic stem cell;  
 KW macrocytic anaemia; aplastic anaemia; cancer; radiation; mouse.  
 XX  
 OS Homo sapiens.  
 OS  
 FH Key Location/Qualifiers  
 FT Peptide 1..27  
 FT Protein /note= "Signal peptide"  
 FT /note= "Mature flk-2"  
 XX US5621090-A.  
 XX 15-APR-1997.  
 XX 02-APR-1991; 91US-0679666.  
 XX 26-JUN-1992; 92US-0906397.  
 XX 02-APR-1991; 91US-0679666.  
 XX 28-JUN-1991; 91US-0728913.  
 XX 15-NOV-1991; 91US-0793065.  
 XX 24-DEC-1991; 91US-0813593.  
 XX (UYPR-) UNIV PRINCETON.  
 XX Lemischka IR;  
 XX WPI; 1997-235228/21.  
 XX N-PSDB; AAT72117.  
 XX Protein containing the extracellular domain of human flk-2 - used  
 XX for identification of primitive haematopoietic cell proliferation  
 XX and differentiation stimulatory ligands, e.g. for treating anaemia  
 XX .Claim 1; Fig 1B; 55pp; English.  
 XX  
 CC This sequence represents human fetal liver kinase 2 (flk2). flk-2 is  
 CC a receptor protein tyrosine kinase (ptk) and is important in transducing  
 CC putative self-renewal signals from the environment. flk-2 is expressed  
 CC in fetal liver, spleen, fetal thymus, adult brain and adult bone marrow,  
 CC and it is thought that flk-2 is expressed in the entire primitive portion  
 CC of the haematopoietic hierarchy. The invention concerns a recombinant  
 CC nucleic acid, preferably mRNA, which encodes a protein containing only  
 CC the extracellular domain of human flk-2 and lacking the flk-2 intra-  
 CC cellular catalytic domain. The resultant protein represents a soluble  
 CC form of flk-2 which is used to isolate specific ligands for flk-2. These  
 CC ligands can be used to stimulate proliferation and/or differentiation of

CC mammalian haematopoietic stem cells, in vivo or in vitro, e.g. for  
CC treatment of macrocytic or aplastic anaemia or bone marrow damage caused  
CC by cancer treatment or radiation.  
XX  
SQ Sequence 993 AA;

Query Match 100.0%; Score 5274; DB 18; Length 993;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVFSAMIFGTITNODLPVVKVNLHNKNDSSVGKSSSPMWSESP 60  
DB 1 MPALARDAGTVPLLVFSAMIFGTITNODLPVVKVNLHNKNDSSVGKSSSPMWSESP 60

QY 61 EDLGCALRPQSSGTVEAAVEVDVSASITLQVLVDAPGNISCLWVFKHSLNCPHFDL 120  
DB 61 EDLGCALRPQSSGTVEAAVEVDVSASITLQVLVDAPGNISCLWVFKHSLNCPHFDL 120

QY 121 QNRGVVSMVILKMTETQAGEYLLFIOSEATNYTLFTVSIRNTLLYTLRPRYFRKMNQD 180  
DB 121 QNRGVVSMVILKMTETQAGEYLLFIOSEATNYTLFTVSIRNTLLYTLRPRYFRKMNQD 180

QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
DB 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240

QY 241 CTLEFTIDLNQTPQTLPOLFLKVGPELWIRCKAVHVNHGFGTLWLENKALEEGNFEM 300  
DB 241 CTLEFTIDLNQTPQTLPOLFLKVGPELWIRCKAVHVNHGFGTLWLENKALEEGNFEM 300

QY 301 STYSTNRTRIRILFAFVSSVARNDTYTCTSSSKHPSQSALVTIVGKGFNATNSSDEYE 360  
DB 301 STYSTNRTRIRILFAFVSSVARNDTYTCTSSSKHPSQSALVTIVGKGFNATNSSDEYE 360

QY 361 IDOYEEFCFSVRKAYPQIRCTWTFKSPCKQGLDNGYSISKFNHKKHQGEYIFHA 420  
DB 361 IDOYEEFCFSVRKAYPQIRCTWTFKSPCKQGLDNGYSISKFNHKKHQGEYIFHA 420

QY 421 ENDDAQTFTKMTLNIRKPOVLAESAQSACFSDGVPPLSWTKKCDKSPNCTEITE 480  
DB 421 ENDDAQTFTKMTLNIRKPOVLAESAQSACFSDGVPPLSWTKKCDKSPNCTEITE 480

QY 481 GVMNRKANRKFQWVSSSTLNNSAIGFLVKCCAYNSLGTSCETILLNSPGFPFIQD 540  
DB 481 GVMNRKANRKFQWVSSSTLNNSAIGFLVKCCAYNSLGTSCETILLNSPGFPFIQD 540

QY 541 NISFYATIGVCLLFIVVLLILCHYKKQPRYESQLQMVQVGTSSDNEYFYDFREYED 600  
DB 541 NISFYATIGVCLLFIVVLLILCHYKKQPRYESQLQMVQVGTSSDNEYFYDFREYED 600

QY 601 LKWEFPRENLEFGKVLGSAFGKVMATAGISKTGVSIOVAVKMLKEKADSSEREALMS 660  
DB 601 LKWEFPRENLEFGKVLGSAFGKVMATAGISKTGVSIOVAVKMLKEKADSSEREALMS 660

QY 661 ELKMMTQOLGSHENIVNLGACTLSGPTLYLFECYCGDNLNLSRKREKFRHTWTEIFKE 720  
DB 661 ELKMMTQOLGSHENIVNLGACTLSGPTLYLFECYCGDNLNLSRKREKFRHTWTEIFKE 720

QY 721 HNFSTYPTFQSHPNMSPGSRVQIHPDSDQISGLHGNFSHSEDEIYENQKRLSEEDL 780  
DB 721 HNFSTYPTFQSHPNMSPGSRVQIHPDSDQISGLHGNFSHSEDEIYENQKRLSEEDL 780

QY 781 NVLTFEDLLCFAYQAKGMEFLFKSCVHRDLAARNVLTGHKVKVICDGLARDINSDS 840  
DB 781 NVLTFEDLLCFAYQAKGMEFLFKSCVHRDLAARNVLTGHKVKVICDGLARDINSDS 840

QY 841 NVVVRNARLPVKWMAPESELFEGIYITKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900  
DB 841 NVVVRNARLPVKWMAPESELFEGIYITKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900

QY 901 LQIONGFKMDQPFYATEEYIYIMOSWAFDSKRKPSFPNLTSLFGCCQLADAEAMYQNVDG 960  
DB 901 LQIONGFKMDQPFYATEEYIYIMOSWAFDSKRKPSFPNLTSLFGCCQLADAEAMYQNVDG 960

QY 961 RVSECPHTYQNRPFESREMDLGLLSPOAQVEDS 993  
DB 961 RVSECPHTYQNRPFESREMDLGLLSPOAQVEDS 993

RESULT 5  
AAAY08617  
ID AAY08617 standard; Protein; 993 AA.  
XX AAY08617;  
AC AAY08617;  
XX  
DT 05-AUG-1999 (first entry)  
XX  
DE Human flk-2 protein.  
XX  
KW Human; flk-2; flk-1; cell isolation; fetal liver kinase; receptor;  
KW monoclonal; polyclonal; antibody; tyrosine kinase.  
XX  
OS Homo sapiens.  
XX  
PN US5912133-A.  
XX  
PD 15-JUN-1999.  
XX  
PF 10-FEB-1998; 98US-0021324.  
XX  
PR 19-NOV-1992; 92US-0977451.  
PR 02-APR-1991; 91US-0679666.  
PR 28-JUN-1991; 91US-0728913.  
PR 15-NOV-1991; 91US-0793065.  
PR 24-DEC-1991; 91US-0813593.  
PR 26-JUN-1992; 92US-0906397.  
PR 12-NOV-1992; 92US-0975049.  
PR 30-APR-1993; 93US-0055269.  
PR 31-OCT-1994; 94US-0252498.  
PR 15-FEB-1996; 96US-0601891.  
XX  
PA (UYPR-) UNIV PRINCETON.  
XX  
PI Lemischka IR;  
XX  
XX WPI; 1999-357194/30.  
XX N-PSDB; AAX77515.  
XX  
PT Isolating hematopoietic cells expressing fetal liver kinase 1  
PT receptors  
XX  
PS Disclosure; Fig 1b; 59pp; English.  
XX  
CC This invention describes a novel method of isolating cells expressing  
CC fetal liver kinase 1 (flk-1) receptors on their surface and comprises  
CC binding the cells to a polyclonal or monoclonal antibody specific to  
CC the flk-1 receptor and isolating the cells that have bound to the  
CC antibody. The method can be used to isolate hematopoietic stem cells in  
CC any mammal but preferably a rat, mouse, rabbit or human. The proteins of  
CC the invention belong to the receptor protein family. This sequence  
CC represents the human flk-2 protein which is used in the method of the  
CC invention.  
XX  
SQ Sequence 993 AA;

Query Match 100.0%; Score 5274; DB 20; Length 993;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVFSAMIFGTITNODLPVVKVNLHNKNDSSVGKSSSPMWSESP 60  
DB 1 MPALARDAGTVPLLVFSAMIFGTITNODLPVVKVNLHNKNDSSVGKSSSPMWSESP 60

QY 61 EDLGCALRPQSSGTVEAAVEVDVSASITLQVLVDAPGNISCLWVFKHSLNCPHFDL 120  
DB 61 EDLGCALRPQSSGTVEAAVEVDVSASITLQVLVDAPGNISCLWVFKHSLNCPHFDL 120



```
QY 121 QNRGVVSMVLKMTETQAGEYLLFIOSEATNYTLFTVSIRNTLLYTLRRPYFRKMENQD 180
DB 121 QNRGVVSMVLKMTETQAGEYLLFIOSEATNYTLFTVSIRNTLLYTLRRPYFRKMENQD 180
QY 181 ALVCISESPEPIVWVLCDSQGESCKEESPVAVVKEEKVHLHFLGFTDIRCCARNELGRE 240
DB 181 ALVCISESPEPIVWVLCDSQGESCKEESPVAVVKEEKVHLHFLGFTDIRCCARNELGRE 240
QY 241 CTRFLTIDLNOTPQTLPQLFLKVGPELWTRCAVHVNHGFGLTWLENKALEEGNYFEM 300
DB 241 CTRFLTIDLNOTPQTLPQLFLKVGPELWTRCAVHVNHGFGLTWLENKALEEGNYFEM 300
QY 301 STYSTNRTMIRILFAFVSSVARNDGTGYTCSSSKHPSQSALVTIVGKGFNATNSSDYE 360
DB 301 STYSTNRTMIRILFAFVSSVARNDGTGYTCSSSKHPSQSALVTIVGKGFNATNSSDYE 360
QY 361 IDOYEFCFSVREKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNHHKHQPGXEYIFHA 420
DB 361 IDOYEFCFSVREKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNHHKHQPGXEYIFHA 420
QY 421 ENDDAFTKMTFLNIRRKPOVLAEASQASQCFSDGYPLPSWTWKKCDKSPNCTEITE 480
DB 421 ENDDAFTKMTFLNIRRKPOVLAEASQASQCFSDGYPLPSWTWKKCDKSPNCTEITE 480
QY 481 GWNRRKANRVFGQWSSSTLNSEATKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
DB 481 GWNRRKANRVFGQWSSSTLNSEATKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
QY 541 NISFYATIGVCLLFIIVVTLILCHYKKQPRYESQLQMVQVGTSSDNEYFYVDFREYED 600
DB 541 NISFYATIGVCLLFIIVVTLILCHYKKQPRYESQLQMVQVGTSSDNEYFYVDFREYED 600
QY 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISTGTGYSIQVAVKMLKEKADSSREALMS 660
DB 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISTGTGYSIQVAVKMLKEKADSSREALMS 660
QY 661 ELKMMTQLGSHENIVNLLGACTLSGPYILIFYCCYGDLLNLYRSKREKFRHTWTETPK 720
DB 661 ELKMMTQLGSHENIVNLLGACTLSGPYILIFYCCYGDLLNLYRSKREKFRHTWTETPK 720
QY 721 HNFSEYPTFQSHPNSSMPGSRVQIHPDSDQISGLHGNFSHSEDEIEYENOKRLEEEEDL 780
DB 721 HNFSEYPTFQSHPNSSMPGSRVQIHPDSDQISGLHGNFSHSEDEIEYENOKRLEEEEDL 780
QY 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840
DB 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840
QY 841 NYVVRGNARLPVKWMAPESLFEGITTKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
DB 841 NYVVRGNARLPVKWMAPESLFEGITTKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900
QY 901 LIQNGFKMDQPFYATEBIYIIMOSCAFDSKRPSPFNLTSLFGLCOLADAEAMQYQNVG 960
DB 901 LIQNGFKMDQPFYATEBIYIIMOSCAFDSKRPSPFNLTSLFGLCOLADAEAMQYQNVG 960
QY 961 RVSECPHTYQNRPRFSEMDLGLLSPQAQVEDS 993
DB 961 RVSECPHTYQNRPRFSEMDLGLLSPQAQVEDS 993

RESULT 6
AAR37503
ID AAR37503 standard; Protein; 993 AA.
XX
AC AAR37503;
XX
DT 19-OCT-1993 (first entry)
XX
DE Human flk-2.
XX
KW Murine; receptor; protein; tyrosine kinase; ptk; flk-2; primitive;
```

hematopoietic cell; mature; family; conserved; region;  
catalytic domain; c-kit; fetal liver kinase; flk; fetal; spleen;  
thymus; adult; brain; bone marrow; multipotential; CFU-Blast colony;  
hierarchy; transduction; T-lymphoid; lineage.

Homo sapiens.

Key Location/Qualifiers  
Peptide 1..27  
Protein /note= "Hydrophobic leader sequence"  
Domain 28..993  
Region /note= "Mature murine flk-2"  
Domain 28..543  
Region /note= "Extracellular receptor domain"  
Domain 544..563  
Domain /note= "Transmembrane region"  
Domain 564..983  
Domain /note= "Intracellular catalytic domain"

W09310136-A.

27-MAY-1993.

16-NOV-1992; 92WO-US09893.

15-NOV-1991; 91US-0793065.

(UYPR-) UNIV PRINCETON.

Lemischka IR;

WPI; 1993-182479/22.

N-PSDB; AAQ40915.

Totipotent haematopoietic stem cell receptors, their ligands and  
DNA sequences - for treating anaemia(s) and bone marrow damage  
due to e.g. cancer chemotherapy or radiotherapy

Claim 39; Fig 1b; 127pp; English.

This sequence represents the human receptor protein tyrosine kinase  
(ptk), flk-2. The nucleic acid encoding this receptor is expressed  
in primitive hematopoietic cells and not in mature hematopoietic  
cells. Members of this family of ptk's can be recognised by the  
conserved amino acid regions in the catalytic domain. This family  
of ptk's also contains c-kit. These new receptors are termed fetal  
liver kinases (flk's) after the tissue in which they were discovered.  
flk-2 is also expressed in fetal spleen, fetal thymus, adult brain  
and adult bone marrow. flk-2 is expressed in individual multipotential  
CFU-Blast colonies capable of generating numerous multilineage colonies  
upon replating. It is likely therefore, that flk-2 is expressed in  
the entire primitive portion of the hematopoietic hierarchy. This is  
consistent with flk-2 being important in transducing putative self-  
renewal signals from the environment. flk-2 is the first receptor  
ptk known to be expressed in the T-lymphoid lineage.

Sequence 993 AA;

Query Match 99.9%; Score 5271; DB 14; Length 993;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 992; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAAGTVPLLVVFSAMIFGTITNODLPVICKVLINHNKNDSSVGKSSYPWVSESP 60

DB 1 MPALARDAAGTVPLLVVFSAMIFGTITNODLPVICKVLINHNKNDSSVGKSSYPWVSESP 60

QY 61 EDLGCALRPQSSGTVYEAAGVVDVSASITLQVLVDAPGNISCLWVFKHSSLNCOPHEDL 120

DB 61 EDLGCALRPQSSGTVYEAAGVVDVSASITLQVLVDAPGNISCLWVFKHSSLNCOPHEDL 120

QY 121 QNRGVVSMVLKMTETQAGEYLLFTQSEATNYTLFTVSIRNTLLYTLRRPYFRKMENQD 180

DB 121 QNRGVVSMVLKMTETQAGEYLLFTQSEATNYTLFTVSIRNTLLYTLRRPYFRKMENQD 180

```
QY 181 ALVCISVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
DB 181 ALVCISVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
QY 241 CTRFLTIDLNOTPQTTLPLQFLKVGEPWIRKAVHVNHGFGLTWELENKALEEGNYFEM 300
DB 241 CTRFLTIDLNOTPQTTLPLQFLKVGEPWIRKAVHVNHGFGLTWELENKALEEGNYFEM 300
QY 301 STYSTNRTMIRILFAFVSSVARNDTGYITCSSKHPQSQALVTIVGKGFNATNSSDEYE 360
DB 301 STYSTNRTMIRILFAFVSSVARNDTGYITCSSKHPQSQALVTIVGKGFNATNSSDEYE 360
QY 361 IDOYEFCFSVREKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHKHPGEYIFHA 420
DB 361 IDOYEFCFSVREKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHKHPGEYIFHA 420
QY 421 ENDDAQFTKMTLNIRRKPOVLAEASQASCFSDGYPLPSWTWKKCDKSPNCTEITE 480
DB 421 ENDDAQFTKMTLNIRRKPOVLAEASQASCFSDGYPLPSWTWKKCDKSPNCTEITE 480
QY 481 GWNRRKANRVFQWVSSSTLNSEAIKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
DB 481 GWNRRKANRVFQWVSSSTLNSEAIKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
QY 541 NISFYATIGVCLLFIVVLTLLCHKYKQPRYESOLOMVQVGTSSDNEYFYVDFREYED 600
DB 541 NISFYATIGVCLLFIVVLTLLCHKYKQPRYESOLOMVQVGTSSDNEYFYVDFREYED 600
QY 601 LKWEFPRENLEFGKVLGSGAFGVMNATAYGISKTGVSIOVAVKMLKEKADSSEREALMS 660
DB 601 LKWEFPRENLEFGKVLGSGAFGVMNATAYGISKTGVSIOVAVKMLKEKADSSEREALMS 660
QY 661 ELKMTQOLASHENIVNLGACTLSGGIYILFECYCCYGDLLNLYRSRKEKPHRTWTEIFKE 720
DB 661 ELKMTQOLASHENIVNLGACTLSGGIYILFECYCCYGDLLNLYRSRKEKPHRTWTEIFKE 720
QY 721 HNFSPYPTQSHNSMSPGSRREVQIHPDSDQISGLHGNSFHSDEIEYENOKRLEEEEDL 780
DB 721 HNFSPYPTQSHNSMSPGSRREVQIHPDSDQISGLHGNSFHSDEIEYENOKRLEEEEDL 780
QY 781 NVLTFEDLLCFAYQVAKGMEFFLFCVHRDLAARNVLTGHGVKVKICDFGLARDIMSDS 840
DB 781 NVLTFEDLLCFAYQVAKGMEFFLFCVHRDLAARNVLTGHGVKVKICDFGLARDIMSDS 840
QY 841 NVVVRGNARLPVKWMAPESLFEGIYITKSDVWSYGILLNWFISLGVNYPGIPVDANFYK 900
DB 841 NVVVRGNARLPVKWMAPESLFEGIYITKSDVWSYGILLNWFISLGVNYPGIPVDANFYK 900
QY 901 LIQNGFKMDQPFYATBEIYIIMOSWAFDSRKRPSPNLTSLFLGCQLADAEAMYQNVDG 960
DB 901 LIQNGFKMDQPFYATBEIYIIMOSWAFDSRKRPSPNLTSLFLGCQLADAEAMYQNVDG 960
QY 961 RVSECPHTYQNRPRFSDREMDLGLLSPOAIVEDS 993
DB 961 RVSECPHTYQNRPRFSDREMDLGLLSPOAIVEDS 993
RESULT 7
ID AAR44995
XX AAR44995 standard; Protein; 993 AA.
AC AAR44995;
XX 27-JUN-1994 (first entry)
XX Human flk-2 receptor protein tyrosine kinase.
XX Receptor protein tyrosine kinase; pTK family; foetal liver kinase;
KW flk; primitive; totipotent; haematopoietic cell; stem cell;
KW proliferation; stromal cell.
XX Homo sapiens.
OS
```

```
XX Key Location/Qualifiers
PH 1..27 /label= signal_sequence
FT /note= "hydrophobic leader"
FT 28..993 /label= flk-2
FT /note= "mature protein"
FT 28..544 /label= extracellular_domain
FT 545..563 /label= transmembrane_region
FT 564..993 /label= intracellular_catalytic_domain
XX
XX US270458-A.
XX 14-DEC-1993.
XX 02-APR-1991; 91US-0679666.
XX 02-APR-1991; 91US-0679666.
XX 28-JUN-1991; 91US-0728913.
XX 15-NOV-1991; 91US-0793065.
XX 24-DEC-1991; 91US-0813593.
XX 26-JUN-1992; 92US-0906397.
XX 12-NOV-1992; 92US-0975049.
XX 19-NOV-1992; 92US-0977451.
XX (UUPR-) UNIV PRINCETON.
XX
XX Lemischka IR;
XX
XX WPI; 1993-405021/50.
XX N-PSDB; AAQ53503.
XX
XX Isolated nucleic acid molecules of hematopoietic stem cell
XX receptor flk-2 - encoding mammalian receptor protein tyrosine
XX kinases expressed in primitive haematopoietic cells
XX
XX Disclosure; Fig 1b; 60pp; English.
XX
XX Nucleic acid sequences coding for murine flk-2 and specified
XX subfragments of it are claimed. The human flk-2 coding sequence
XX (i.e. AAQ53503) is also disclosed. The murine and human flk-2
XX polypeptides represent a new class of receptor protein tyrosine
XX kinases which are expressed only in primitive haematopoietic cells.
XX
XX Sequence 993 AA;
XX
XX Query Match 99.8%; Score 5266; DB 14; Length 993;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 992; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MPALARDAGTVPPLLVVFSAMIFGTITNODLPVIRKVLINHKNDSSVCKSSYPWVSESP 60
DB 1 MPALARDAGTVPPLLVVFSAMIFGTITNODLPVIRKVLINHKNDSSVCKSSYPWVSESP 60
QY 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVKHSSLNCPHFDL 120
DB 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVKHSSLNCPHFDL 120
QY 121 QNRGVVSMVILKMTETQAGEYLLFIQSEATNYTILFTVSIIRNTLLYTLRRPYFRKMNQD 180
DB 121 QNRGVVSMVILKMTETQAGEYLLFIQSEATNYTILFTVSIIRNTLLYTLRRPYFRKMNQD 180
QY 181 ALVCISVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
DB 181 ALVCISVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240
QY 241 CTRFLTIDLNOTPQTTLPLQFLKVGEPWIRKAVHVNHGFGLTWELENKALEEGNYFEM 300
DB 241 CTRFLTIDLNOTPQTTLPLQFLKVGEPWIRKAVHVNHGFGLTWELENKALEEGNYFEM 300
```

QY	301	STYSTNRTMIRILFAFVSSVARNDTGYYTCSSSKHPSQSALVTIVGKGFNATNSSDYE	360
Db	301	STYSTNRTMIRILFAFVSSVARNDTGYYTCSSSKHPSQSALVTIVGKGFNATNSSDYE	360
QY	361	IDQEEFCFSVRFAKYPQIRCTWTFSRKFPCEOKGLDNGYSISKFCNKHQHPGEYIFHA	420
Db	361	IDQEEFCFSVRFAKYPQIRCTWTFSRKFPCEOKGLDNGYSISKFCNKHQHPGEYIFHA	420
QY	421	ENDDAQFTKMTLNIRRKPOVLAESAQSCFSDGYPLPSWTWKKCSKSPNCTEITE	480
Db	421	ENDDAQFTKMTLNIRRKPOVLAESAQSCFSDGYPLPSWTWKKCSKSPNCTEITE	480
QY	481	GWNRKANKRVGQWVSSSTLNNSAIAKFLVKCCAYNSLGTSCETILLNSPGPFPIQD	540
Db	481	GWNRKANKRVGQWVSSSTLNNSAIAKFLVKCCAYNSLGTSCETILLNSAGFFPIQD	540
QY	541	NTSFYATIGVCLLFIIVVLTLLCHKYYKKQFRYESQIQVAVKMLKEKADSSEREA	600
Db	541	NTSFYATIGVCLLFIIVVLTLLCHKYYKKQFRYESQIQVAVKMLKEKADSSEREA	600
QY	601	LKWEPPRENLEFGKVLGSAFGKVMNATAYGISKTVGSIQVAVKMLKEKADSSEREA	660
Db	601	LKWEPPRENLEFGKVLGSAFGKVMNATAYGISKTVGSIQVAVKMLKEKADSSEREA	660
QY	661	ELKMMTQLGSHENIVNLLGACTLSGPIYLIFFBYCCYGDLLNLYLSRKREKPHRTWT	720
Db	661	ELKMMTQLGSHENIVNLLGACTLSGPIYLIFFBYCCYGDLLNLYLSRKREKPHRTWT	720
QY	721	HNFSFYPTQSHNSMPGSRREVQIHPDSDQISGLHGNPSHSEDETEYENOKRLEED	780
Db	721	HNFSFYPTQSHNSMPGSRREVQIHPDSDQISGLHGNPSHSEDETEYENOKRLEED	780
QY	781	NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTHTGKVKVVICDFGLARD	840
Db	781	NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTHTGKVKVVICDFGLARD	840
QY	841	NYVVRGNARLPVKWMAPESLFEGIIYIKSDVWSYGILLWEIFSLGVNPNYPGIPV	900
Db	841	NYVVRGNARLPVKWMAPESLFEGIIYIKSDVWSYGILLWEIFSLGVNPNYPGIPV	900
QY	901	LIONGFKMDOPFYATEIYIIMOSWAFDSRKRPSFNLTSFLGCOLADAEAMYNQV	960
Db	901	LIONGFKMDOPFYATEIYIIMOSWAFDSRKRPSFNLTSFLGCOLADAEAMYNQV	960
QY	961	RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS	993
Db	961	RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS	993
RESULT 8			
AAR75961			
ID	AAR75961 standard; Protein: 993 AA.		
XX	AC AAR75961;		
XX	29-DEC-1995 (first entry)		
DT	Human STK-1.		
DE	STK-1; receptor PTK; protein tyrosine kinase.		
KW	Homo sapiens.		
XX	Key Location/Qualifiers		
XX	Peptide 1..23		
FT	/label= signal		
FT	/note= "unique PTK receptor sequence"		
FT	Cleavage-site 35..226		
FT	Domain 25..524		
FT	/note= "Part of extracellular domain contg. 22 Cys residues marked on Fig. 1"		
FT	Modified-site 43..45		

FT	Modified-site	/label= potential N-linked glycosylation site
FT	Modified-site	100..102
FT	Modified-site	/label= see above
FT	Modified-site	151..153
FT	Modified-site	/label= see above
FT	Modified-site	250..252
FT	Modified-site	/label= see above
FT	Modified-site	306..308
FT	Modified-site	/label= see above
FT	Modified-site	323..325
FT	Modified-site	/label= see above
FT	Modified-site	351..353
FT	Modified-site	/label= see above
FT	Modified-site	354..356
FT	Modified-site	/label= see above
FT	Modified-site	473..475
FT	Modified-site	/label= see above
FT	Modified-site	502..504
FT	Modified-site	/label= see above
FT	Region	542..562
FT	Region	/label= transmembrane spanning region
FT	Region	708..782
FT	Region	/label= kinase insert region
FT	Domain	617..622
FT	Domain	/label= ATP binding domain
FT	Domain	835..840
FT	Domain	/label= WMAPES motifs
FT	Domain	/note= "cytoplasmic domain"
FT	Peptide	808..813
FT	Peptide	/note= "used to design PCR oligos"
FT	Peptide	870..875
FT	Peptide	/note= "used to design PCR oligos"
XX	W09519175-A.	
PN	20-JUL-1995.	
XX	06-JAN-1995; 95WO-US00176.	
XX	14-JAN-1994; 94US-0183211.	
PA	(UYJO ) UNIV JOHNS HOPKINS.	
PA	(UYPE-) UNIV PENNSYLVANIA.	
XX	Civin CI, Gewirtz AM, Small D;	
XX	WPI; 1995-263709/34.	
DR	N-PSDB; AAQ91356.	
XX	Artificial STK-1 gene and gene-specific anti-sense oligo:nucleotide -	
PT	used to treat neoplastic diseases and as bone marrow purging agents for	
PT	treating leukaemia and neoplasma	
XX	Disclosure; Fig 1; 66pp; English.	
XX	The STK-1 gene encodes a receptor PTK which is expressed in	
CC	proliferating hematopoietic stem cells but not in quiescent stem	
CC	cells. The STK-1 gene is also expressed in certain malignant cells	
CC	of non-hematopoietic origin. An antisense oligo specific for STK-1	
CC	is an oligo having a sequence (i) capable of forming a stable	
CC	triplex with a portion of the STK-1 gene, or (ii) capable of forming	
CC	a stable duplex with a portion of an mRNA transcript of the STK-1	
CC	gene. Antisense oligos capable of forming a stable duplex with a	
CC	portion of a STK-1 mRNA transcript are given in AAQ91536 FT and in	
CC	AAQ91537 and AAQ91538. The antisense oligos of the invention are useful	
CC	in the treatment of hematologic malignancies characterised by	
CC	STK-1 expression. Several of the conserved domains of PTKs	
CC	including the ATP binding domain and the WMAPES motifs are	
CC	found in the STK-1 protein (see AAR75961 FT).	
XX	Sequence 993 AA;	
SQ	Query Match 99.8%; Score 5266; DB 16; Length 993;	

Best Local Similarity 99.9%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVVFSAAMIFGTITNODLPVVKVCLINHKNDSSVSKSSYPMVSESP 60  
DB 1 MPALARDAGTVPLLVVFSAAMIFGTITNODLPVVKVCLINHKNDSSVSKSSYPMVSESP 60  
QY 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSLNCOPHFDL 120  
DB 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSLNCOPHFDL 120  
QY 121 QNRGVVSMVLKMTETQAGEYLLFIOSEATNYTLFTVSVIRNTLLYLRPPYFRKMENQD 180  
DB 121 QNRGVVSMVLKMTETQAGEYLLFIOSEATNYTLFTVSVIRNTLLYLRPPYFRKMENQD 180  
QY 181 ALVCISESPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
DB 181 ALVCISESPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
QY 241 CTRFLTIDLNOTPTQTLPOLFLKVGEPWIRCKAVHVNHGFGLTWELENKALEGNYFEM 300  
DB 241 CTRFLTIDLNOTPTQTLPOLFLKVGEPWIRCKAVHVNHGFGLTWELENKALEGNYFEM 300  
QY 301 STYSTNRTMIRILFAFVSSVARNDGTGYTCCSSKHPQSQSALVTIVGKGFINATNSSDEYE 360  
DB 301 STYSTNRTMIRILFAFVSSVARNDGTGYTCCSSKHPQSQSALVTIVGKGFINATNSSDEYE 360  
QY 361 IDQYEFECFSVRKAYPOIRCTWTFRSKSPCEQKGLDNGYSISKFCNHRKHQGEYIFHA 420  
DB 361 IDQYEFECFSVRKAYPOIRCTWTFRSKSPCEQKGLDNGYSISKFCNHRKHQGEYIFHA 420  
QY 421 ENDDAOFKMTINIRKPOVLAEASASQSCFSDGYPPLSWTKKCDKSPNCTEITE 480  
DB 421 ENDDAOFKMTINIRKPOVLAEASASQSCFSDGYPPLSWTKKCDKSPNCTEITE 480  
QY 481 GVMNRKANRKFQWVSSSTLNHSEAIKGLVKCCAYNSLGTSCETILLNSPGFPFIQD 540  
DB 481 GVMNRKANRKFQWVSSSTLNHSEAIKGLVKCCAYNSLGTSCETILLNSPGFPFIQD 540  
QY 541 NTSFYATIGVCLLFIYVLLILCHIKYKQFRYESQLOMVQVTSDDNEYYVDFREYEYD 600  
DB 541 NTSFYATIGVCLLFIYVLLILCHIKYKQFRYESQLOMVQVTSDDNEYYVDFREYEYD 600  
QY 601 LKWEFPRENLEFGKVLGSGAFGVMNATAYGISTGVSIOVAVKMLKEKADSSERELMS 660  
DB 601 LKWEFPRENLEFGKVLGSGAFGVMNATAYGISTGVSIOVAVKMLKEKADSSERELMS 660  
QY 661 ELKMMTQLGSHENIVNLGACTLSGPIYLFYCYCYGDLNLYRSRKEKPHRTWTLEFKE 720  
DB 661 ELKMMTQLGSHENIVNLGACTLSGPIYLFYCYCYGDLNLYRSRKEKPHRTWTLEFKE 720  
QY 721 HNFSEYPTQSHPNSSMPGSRVQIHPDSQIISGLHGNPSHSEDETEYENQKRLSEEDL 780  
DB 721 HNFSEYPTQSHPNSSMPGSRVQIHPDSQIISGLHGNPSHSEDETEYENQKRLSEEDL 780  
QY 781 NVLTFEDLLCFAYQAVAKMEFLFKSCVHRDLAARNVLTGHGVKICDPLGARDIMSDS 840  
DB 781 NVLTFEDLLCFAYQAVAKMEFLFKSCVHRDLAARNVLTGHGVKICDPLGARDIMSDS 840  
QY 841 NVVVRGNARLPVKWMAPELSFEGIYTIKSDVWSYGILLMEIFSLGVPYPCIPVDANFYK 900  
DB 841 NVVVRGNARLPVKWMAPELSFEGIYTIKSDVWSYGILLMEIFSLGVPYPCIPVDANFYK 900  
QY 901 LQNGFKMDQPPYATEEIIYIMQSWAFPSRKRPSFNLTSLFGCOLADAEAMYNQVGD 960  
DB 901 LQNGFKMDQPPYATEEIIYIMQSWAFPSRKRPSFNLTSLFGCOLADAEAMYNQVGD 960  
QY 961 RVSECPHTYQNRPPRFREMDLGLLSPOAQVEDS 993  
DB 961 RVSECPHTYQNRPPRFREMDLGLLSPOAQVEDS 993

RESULT 9

AAR81869  
ID AAR81869 standard; Protein; 993 AA.  
XX  
AC AAR81869;  
XX  
DT 19-MAR-1996 (first entry)  
XX  
DE Human Flk2/flt3 tyrosine kinase receptor.  
XX  
KW Tyrosine kinase receptor; flk2; flt3; agonist; monoclonal antibody;  
KW haematopoiesis; hypoplasia; anaemia; thrombocytopenia; stem cell.  
XX  
OS Homo sapiens.  
XX  
PN W09527062-A1.  
XX  
PD 12-OCT-1995.  
XX  
PF 23-MAR-1995; 95WO-US03718.  
XX  
PR 04-APR-1994; 94US-0222299.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Bennett BD, Broz SD, Matthews W, Zeigler FC;  
XX  
DR WPI; 1995-358636/46.  
DR N-PSDB; AAT00802.  
XX  
PT Agonist antibody which binds to flk2-flt3 tyrosine kinase receptor  
PT - enhances proliferation of haematopoietic stem cells, in the  
PT treatment of hypoplasia, anaemia, etc.  
XX  
PS Disclosure; Page 44-47; 59pp; English.  
XX  
CC DNA (AAT00801) coding for murine stem cell tyrosine kinase receptor flk2  
CC (also called flt3) (AAR81868) was obtd. by RT-PCR amplification of RNA  
CC isolated from mid-gestation mouse foetal livers using primers based  
CC on the murine flt3 sequence, and subcloning of the product into pRK5.1.  
CC An flk2/flt3 extracellular domain-IgG1 Fc fusion protein was  
CC constructed and used to raise agonist antibodies able to bind to, and  
CC activate, flk2/flt3. The human flk2/flt3 receptor gene (AAT00802) and  
CC amino acid sequence (AAR81869) are also given.  
XX  
SQ Sequence 993 AA;  
Query Match 99.8%; Score 5265; DB 16; Length 993;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 992; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVVFSAAMIFGTITNODLPVVKVCLINHKNDSSVSKSSYPMVSESP 60  
DB 1 MPALARDAGTVPLLVVFSAAMIFGTITNODLPVVKVCLINHKNDSSVSKSSYPMVSESP 60  
QY 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSLNCOPHFDL 120  
DB 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSLNCOPHFDL 120  
QY 121 QNRGVVSMVLKMTETQAGEYLLFIOSEATNYTLFTVSVIRNTLLYLRPPYFRKMENQD 180  
DB 121 QNRGVVSMVLKMTETQAGEYLLFIOSEATNYTLFTVSVIRNTLLYLRPPYFRKMENQD 180  
QY 181 ALVCISESPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
DB 181 ALVCISESPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
QY 241 CTRFLTIDLNOTPTQTLPOLFLKVGEPWIRCKAVHVNHGFGLTWELENKALEGNYFEM 300  
DB 241 CTRFLTIDLNOTPTQTLPOLFLKVGEPWIRCKAVHVNHGFGLTWELENKALEGNYFEM 300  
QY 301 STYSTNRTMIRILFAFVSSVARNDGTGYTCCSSKHPQSQSALVTIVGKGFINATNSSDEYE 360  
DB 301 STYSTNRTMIRILFAFVSSVARNDGTGYTCCSSKHPQSQSALVTIVGKGFINATNSSDEYE 360





QY 651 DSSREALMSKMTQLGSHENIVNLGACTLSGPIYLIFEYCCYDGLLNLRSKREKF 710  
 DB 661 DSSREALMSKMTQLGSHENIVNLGACTLSGPIYLIFEYCCYDGLLNLRSKREKF 720  
 QY 711 HRTWTEIFKEHNFSPYPTQSHPNSSMPGSRREVQIHPDSQISGLHGSFHSDEIEIEN 770  
 DB 721 HRTWTEIFKEHNFSPYPTQSHPNSSMPGSRREVQIHPDSQISGLHGSFHSDEIEIEN 780  
 QY 771 OKRLEEEEDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDF 830  
 DB 781 OKRLEEEEDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDF 840  
 QY 831 GLARDIMSDSNVVRGNARLPVKWMAPELSEFEGIYTIKSDVMSYGILLWEIFSLGVPNPYP 890  
 DB 841 GLARDIMSDSNVVRGNARLPVKWMAPELSEFEGIYTIKSDVMSYGILLWEIFSLGVPNPYP 900  
 QY 891 GIPVDANFYKLIQNGFKMDQPFYATEEIIYIMQSCWAFDSRKRPSFPNLTSLGCOLADA 950  
 DB 901 GIPVDANFYKLIQNGFKMDQPFYATEEIIYIMQSCWAFDSRKRPSFPNLTSLGCOLADA 960  
 QY 951 EAMYNQVGRVSECPHTYQNR 973  
 DB 961 EAMYNQVGRVSECPHTYQNR 983

## RESULT 12

AAW63587  
 ID AAW63587 standard; Protein; 986 AA.

AC AAW63587;

DT 12-OCT-1998 (first entry)

DE Human receptor type protein kinase FLT3 protein SEQ ID NO:18.

KW Human; receptor type protein kinase; FLT3; diagnosis; leukaemia;  
 KW membrane-vicinal region; tyrosine kinase; juxtamembrane region.

OS Homo sapiens.

PN WO9817808-A1.

XX 30-APR-1998.

XX 13-OCT-1997; 97WO-JP03667.

XX 18-OCT-1996; 96JP-0297329.

XX (TAKI ) TAKARA SHUZO CO LTD.

XX Yokota S;

XX WPI; 1998-362333/31.

DR N-PSDB; AAV39040.

PT Nucleic acid sequences encoding receptor type protein kinase -  
 PT useful for diagnosing adult acute leukaemia of, e.g. FAB sub-classes  
 PT M2

XX Claim 6; Page 40-45; 80pp; Japanese.

CC New nucleic acid sequences have been isolated which encode receptor type  
 CC protein kinases (especially a tyrosine kinases) having tandem repeats in  
 CC the juxtamembrane region. Also described in the present invention are:  
 CC (1) any one of 10 amino acid (aa) sequences (AAW63580 to AAW63589),  
 CC representing FLT3 juxtamembrane receptor type protein kinases found in  
 CC leukaemia patients of FAB subclass M2, M4 or M5 (nucleic acid sequences  
 CC encoding the proteins are given in AAV39028 to AAV39042); (2) antibodies  
 CC recognising the kinases or their portions including with the tandem repeat  
 CC region; (3) nucleic acid sequences hybridising with the nucleic acids;  
 CC (4) a method for detecting the nucleic acid sequences in human tissue  
 CC samples by: (i) extracting nucleic acid; (ii) amplifying the kinase  
 CC gene, and (iii) determining the size of the gene for comparing the size

CC of the normal gene not containing tandem repeats, and (5) kits for  
 CC carrying out the detection. The products and methods may be used for  
 CC diagnosing adult acute leukaemia of FAB subclasses M2, M4 and M5.

XX Sequence 986 AA;

QY Query Match 97.8%; Score 5157.5; DB 19; Length 986;

DB Best Local Similarity 98.7%; Pred. No. 0;

XX Matches 973; Conservative 0; Mismatches 0; Indels 13; Caps 1;

QY 1 MPALARDAGTVPLLVVSAMIFGTITNODLPVTKCVLINHKNNDSSVGKSSSPYMWSESP 60

DB 1 MPALARDAGTVPLLVVSAMIFGTITNODLPVTKCVLINHKNNDSSVGKSSSPYMWSESP 60

QY 61 EDLGCALRQSSGTVYAAAIVEDVSASITLQVLVDAPGNISCLWFKHSLNCPHFDL 120

DB 61 EDLGCALRQSSGTVYAAAIVEDVSASITLQVLVDAPGNISCLWFKHSLNCPHFDL 120

QY 121 QNRGVVSMVLKMTETQAGEYLLFIOSEATNYTILFTVSRINTLLYTLRRPYFRKMNQD 180

DB 121 QNRGVVSMVLKMTETQAGEYLLFIOSEATNYTILFTVSRINTLLYTLRRPYFRKMNQD 180

QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240

DB 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240

QY 241 CTRLEFTIDLNOTPQTTLPQLFLKVGPELWIRCAVHVNHGFGLTWELENKALEGNGFEM 300

DB 241 CTRLEFTIDLNOTPQTTLPQLFLKVGPELWIRCAVHVNHGFGLTWELENKALEGNGFEM 300

QY 301 STYSTNRTMIRILFAFVSSVARNDTGYYTCTSSSKHPQSALVTIIVGKGFNATNSSEDE 360

DB 301 STYSTNRTMIRILFAFVSSVARNDTGYYTCTSSSKHPQSALVTIIVGKGFNATNSSEDE 360

QY 361 IDQYEFCEFSVRKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFNHKKHQGEYIFHA 420

DB 361 IDQYEFCEFSVRKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFNHKKHQGEYIFHA 420

QY 421 ENDDAOTFKMFTLNIRKPOVLAEASASQSCFSDGYPPLSWTWKCKSDKSPNCTEITE 480

DB 421 ENDDAOTFKMFTLNIRKPOVLAEASASQSCFSDGYPPLSWTWKCKSDKSPNCTEITE 480

QY 481 GVNRRKANRVFGQWSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540

DB 481 GVNRRKANRVFGQWSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540

QY 541 NISFYATIGVCLLFIYVLTLLICHKKYKQFRYESQLQWQVGTGSSDNEYFYV----- 592

DB 541 NISFYATIGVCLLFIYVLTLLICHKKYKQFRYESQLQWQVGTGSSDNEYFYV----- 592

QY 593 -----DPREYEDLKWEEFPRENLEFGKVGSGAGKVMNATAYGISTGVSIOAVKMLK 647

DB 601 LKWEFDREYEDLKWEEFPRENLEFGKVGSGAGKVMNATAYGISTGVSIOAVKMLK 660

QY 648 EKADSSEREALMSKMTQLGSHENIVNLGACTLSGPIYLIFEYCCYDGLLNLRSKR 707

DB 661 EKADSSEREALMSKMTQLGSHENIVNLGACTLSGPIYLIFEYCCYDGLLNLRSKR 720

QY 708 EXPFRTWTEIFKEHNFSPYPTQSHPNSSMPGSRREVQIHPDSQISGLHGSFHSDEIE 767

DB 721 EXPFRTWTEIFKEHNFSPYPTQSHPNSSMPGSRREVQIHPDSQISGLHGSFHSDEIE 780

QY 768 YENQKRLLEEEDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKI 827

DB 781 YENQKRLLEEEDNLVTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKI 840

QY 828 CDFGLARDIMSDSNVVRGNARLPVKWMAPELSEFEGIYTIKSDVMSYGILLWEIFSLGVN 887

DB 841 CDFGLARDIMSDSNVVRGNARLPVKWMAPELSEFEGIYTIKSDVMSYGILLWEIFSLGVN 900

QY 888 PYGIPVDANFYKLIQNGFKMDQPFYATEEIIYIMQSCWAFDSRKRPSFPNLTSLGCOL 947

DB 901 PYGIPVDANFYKLIQNGFKMDQPFYATEEIIYIMQSCWAFDSRKRPSFPNLTSLGCOL 960



Qy 948 ADAEAMYNQVNDGRVSECPHTYQNR 973  
 Db 961 ADAEAMYNQVNDGRVSECPHTYQNR 986

## RESULT 13

AAW63589  
 ID AAW63589 standard; Protein; 986 AA.

XX AC AAW63589;

XX DT 12-OCT-1998 (first entry)

XX DE Human receptor type protein kinase FLT3 protein SEQ ID NO:20.

XX KW Human; receptor type protein kinase; FLT3; diagnosis; leukaemia;  
 XX KM membrane-vicinal region; tyrosine kinase; juxtamembrane region.

XX OS Homo sapiens.

XX PN W09817808-A1.

XX PD 30-APR-1998.

XX PF 13-OCT-1997; 97WO-JF03667.

XX PR 18-OCT-1996; 96JP-0297329.

XX PA (TAKI ) TAKARA SHUZO CO LTD.

XX PI Yokota S;

XX DR WPI: 1998-362333/31.

XX DR N-PSDB; AAV39042.

XX PT Nucleic acid sequences encoding receptor type protein kinase -  
 PT useful for diagnosing adult acute leukaemia of, e.g. FAB sub-classes  
 PT M2

XX PS Claim 6; Page 50-55; 80pp; Japanese.

XX CC New nucleic acid sequences have been isolated which encode receptor type  
 CC protein kinases (especially a tyrosine kinases) having tandem repeats in  
 CC the juxtamembrane region. Also described in the present invention are:  
 CC (1) any one of 10 amino acid (aa) sequences (AAW63580 to AAW63589)  
 CC representing FLT3 juxtamembrane receptor type protein kinases found in  
 CC leukaemia patients of FAB subclass M2, M4 or M5 (nucleic acid sequences  
 CC encoding the proteins are given in AAV39028 to AAV39042); (2) antibodies  
 CC recognising the kinases or their portions including the tandem repeat  
 CC region; (3) nucleic acid sequences hybridising with the nucleic acids;  
 CC (4) a method for detecting the nucleic acid sequences in human tissue  
 CC samples by: (i) extracting nucleic acid; (ii) amplifying the kinase  
 CC gene, and (iii) determining the size of the gene for comparing the size  
 CC of the normal gene not containing tandem repeats; and (5) kits for  
 CC carrying out the detection. The products and methods may be used for  
 CC diagnosing adult acute leukaemia of FAB subclasses M2, M4 and M5.

XX Sequence 986 AA;

Query Match 97.8%; Score 5157.5; DB 19; Length 986;  
 Best Local Similarity 98.7%; Pred. No. 0;  
 Matches 973; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

Qy 1 MPALADAGTVPLLVFVSAMIFGTITNQDLPVTKVLIHKNNDSSVCKSSYPMVSESP 60

Db 1 MPALADAGTVPLLVFVSAMIFGTITNQDLPVTKVLIHKNNDSSVCKSSYPMVSESP 60

Qy 61 EDLGCALRPOSSCTVYEAAREVDVSASITLQVLVDAPGNISCLWFKHSLNCQPHFDL 120

Db 61 EDLGCALRPOSSCTVYEAAREVDVSASITLQVLVDAPGNISCLWFKHSLNCQPHFDL 120

Qy 121 ONRGVSVWILKMTQTQAGEYLLFIOSEATNYTLFTVSRNTLLYLRPFYRKMEQND 180

Db 121 ONRGVSVWILKMTQTQAGEYLLFIOSEATNYTLFTVSRNTLLYLRPFYRKMEQND 180  
 Qy 181 ALVCISESVPEPIVWVLCDSOGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
 Db 181 ALVCISESVPEPIVWVLCDSOGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
 Qy 241 CTRLFITDLNQTPTTLPOLFLKVGEPWIRCKAVHNVHGFGLTWELNKALEEGNYFEM 300  
 Db 241 CTRLFITDLNQTPTTLPOLFLKVGEPWIRCKAVHNVHGFGLTWELNKALEEGNYFEM 300  
 Qy 301 STYSTNRTMIRLFAFVSSVARNDTGYTCSSSKHPQSOSALVTIVGKGFINATNSSDEYE 360  
 Db 301 STYSTNRTMIRLFAFVSSVARNDTGYTCSSSKHPQSOSALVTIVGKGFINATNSSDEYE 360  
 Qy 361 IDQYEEFCFSVRFKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQBPGEYIFHA 420  
 Db 361 IDQYEEFCFSVRFKAYPQIRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQBPGEYIFHA 420  
 Qy 421 ENDDAQFTKMTFLNIRRKPOVLAEASASOACFSOGYPLPSWTWKKCDKSPNCTEITE 480  
 Db 421 ENDDAQFTKMTFLNIRRKPOVLAEASASOACFSOGYPLPSWTWKKCDKSPNCTEITE 480  
 Qy 481 GWNKRKANRVFGQWVSSSTLNNSAIGFLVKCCAYNSLGTSCETILLNSGPPFPFIOD 540  
 Db 481 GWNKRKANRVFGQWVSSSTLNNSAIGFLVKCCAYNSLGTSCETILLNSGPPFPFIOD 540  
 Qy 541 NISFYATIGVCLLFTIVLTLTLLCHYKKQFRYESQLOMVQVT-----GSSDN 587  
 Db 541 NISFYATIGVCLLFTIVLTLTLLCHYKKQFRYESQLOMVQVT-----GSSDN 587  
 Qy 588 EYFYVDFREYEDLKWEFPRENLEFGKVLGSGAFKVMNATAYGISKTVSIOVAVKMLK 647  
 Db 601 EYFYVDFREYEDLKWEFPRENLEFGKVLGSGAFKVMNATAYGISKTVSIOVAVKMLK 660  
 Qy 648 EKADSSEREALMSLKMMTOLGSHENIVNLGACTLSGPIYLIFEYCCYCDLLNLYRSKR 707  
 Db 661 EKADSSEREALMSLKMMTOLGSHENIVNLGACTLSGPIYLIFEYCCYCDLLNLYRSKR 720  
 Qy 708 EKFHRTWTEIFKEHNFSFYPTFQSHPNSSMPGSRREVQIHPDSDQISGLHGNFSHSEDEIE 767  
 Db 721 EKFHRTWTEIFKEHNFSFYPTFQSHPNSSMPGSRREVQIHPDSDQISGLHGNFSHSEDEIE 780  
 Qy 768 YENOKRLEEEEDLNVLTFEDLLCFAYQVAKGMFELEFSCVHRDLAARNVLVTHGKVVKI 827  
 Db 781 YENOKRLEEEEDLNVLTFEDLLCFAYQVAKGMFELEFSCVHRDLAARNVLVTHGKVVKI 840  
 Qy 828 CDFGLARDIMSDSNVVRGNARLPVKWMAPELSLFGIYTIKSDVMSYGILLWEIFSLGVN 887  
 Db 841 CDFGLARDIMSDSNVVRGNARLPVKWMAPELSLFGIYTIKSDVMSYGILLWEIFSLGVN 900  
 Qy 888 PYGIPVDANFYKLIQNGFKMDQPFYATEEIIYIMQSCWAFDSRRKRPFPNLTSLGCOL 947  
 Db 901 PYGIPVDANFYKLIQNGFKMDQPFYATEEIIYIMQSCWAFDSRRKRPFPNLTSLGCOL 960  
 Qy 948 ADAEAMYNQVNDGRVSECPHTYQNR 973  
 Db 961 ADAEAMYNQVNDGRVSECPHTYQNR 986

## RESULT 14

AAW63586

ID AAW63586 standard; Protein; 994 AA.

XX AC AAW63586;

XX DT 12-OCT-1998 (first entry)

XX DE Human receptor type protein kinase FLT3 protein SEQ ID NO:17.

XX KW Human; receptor type protein kinase; FLT3; diagnosis; leukaemia;  
 XX KM membrane-vicinal region; tyrosine kinase; juxtamembrane region.

QY	421	ENDDAQTKMFTLNIRRKQVLAELASASQACSFSDGYPPLPSWTWKKCSDKSPNCTEEITE	480
Db	421	ENDDAQTKMFTLNIRRKQVLAELASASQACSFSDGYPPLPSWTWKKCSDKSPNCTEEITE	480
QY	481	GWNRKARKRVQGWSSSTLNMSEBAKGLVKCAYNLSLGTSCETILLNSPGPPFIQD	540
Db	481	GWNRKARKRVQGWSSSTLNMSEBAKGLVKCAYNLSLGTSCETILLNSPGPPFIQD	540
QY	541	NISFYATIGVCLLFIVVLILLCHYKKQFRYESQLQWQVVTGSSDNEYFYVDFREYD	597
Db	541	NISFYATIGVCLLFIVVLILLCHYKKQFRYESQLQWQVVTGSSDNEYFYVDFREYD	600
QY	598	-----EYDLKWEPRENLEFGKVLGSGAFGKVMNATAYGISKTGVSI	639
Db	601	LKWEFPRENLEFGKVLGSEYDLKWEPRENLEFGKVLGSGAFGKVMNATAYGISKTGVSI	660
QY	640	QVAVKMLKEKADSSERREALMSELKMTQLGSHENIVNLLGACTLSGPIYLIFEYCYGDL	699
Db	661	QVAVKMLKEKADSSERREALMSELKMTQLGSHENIVNLLGACTLSGPIYLIFEYCYGDL	720
QY	700	LNLYSRKREKFKHTWTETIEFKHNFSYPTFQSHPNSSMPGSRREVQIHPDSDQISGLHGS	759
Db	721	LNLYSRKREKFKHTWTETIEFKHNFSYPTFQSHPNSSMPGSRREVQIHPDSDQISGLHGS	780
QY	760	FHSEDEIYEYNQRLKEEEDLNVLTFEDLLCFAYQAKMGFLPKSCVHRDLAARNVLV	819
Db	781	FHSEDEIYEYNQRLKEEEDLNVLTFEDLLCFAYQAKMGFLPKSCVHRDLAARNVLV	840
QY	820	THGKVKVICKDFGLARDIMSDSNVVRGNARLPVKWMAPELSFEGIYTIKSDVMSYGILLW	879
Db	841	THGKVKVICKDFGLARDIMSDSNVVRGNARLPVKWMAPELSFEGIYTIKSDVMSYGILLW	900
QY	880	EIFSGLGNVPYGPVPDANFYKLLQNGFKMDQPPYATEEYIIMQSCWAFDSRKRPSFPNL	939
Db	901	EIFSGLGNVPYGPVPDANFYKLLQNGFKMDQPPYATEEYIIMQSCWAFDSRKRPSFPNL	960
QY	940	TSFLGCOLADAEEMAYQNVDRGVSECPHTTYONRR	973
Db	961	TSFLGCOLADAEEMAYQNVDRGVSECPHTTYONRR	994
RESULT 15			
AAR81868			
ID	AAR81868 standard; Protein; 1000 AA.		
XX	AAR81868;		
XX	AC		
XX	AC		
DT	29-FEB-1996 (first entry)		
XX	Flk2/flt3 tyrosine kinase receptor.		
DE			
XX	Tyrosine kinase receptor; flk2; flt3; agonist; monoclonal antibody;		
KW	haematopoiesis; hypoplasma; anaemia; thrombocytopenia; stem cell.		
XX			
OS	Mus sp.		
XX			
FH	Key		
FT	Domain		
FT	1..542		
FT	/label= Extracellular_domain		
XX	W09527062-A1.		
PN			
XX	12-OCT-1995.		
XX			
XX	23-MAR-1995; 95WO-US03718.		
XX			
XX	04-APR-1994; 94US-0222299.		
XX			
PA	(GETH ) GENENTECH INC.		
XX			
PI	Bennett BD, Broz SD, Matthews W, Zeigler FC;		
DR	WPI; 1995-358636/46.		

DR	N-PSDB; AAT00801.	
XX	Agonist antibody which binds to flk2-flt3 tyrosine kinase receptor	
PT	- enhances proliferation of haematopoietic stem cells, in the	
PT	treatment of hypoplasia, anaemia, etc.	
XX	Example 1; Page 38-40; 59pp; English.	
PS		
XX	Murine stem cell tyrosine kinase receptor flk2 (also called flt3)	
CC	(AAR81869) is encoded DNA (AAT00801) obtd. by RT-PCR amplification of	
CC	RNA isolated from mid-gestation mouse foetal livers. An flk2/flt3	
CC	extracellular domain-IgG1 Fc fusion protein was used to raise	
CC	agonist antibodies able to bind to, and activate, flk2/flt3.	
CC	Such antibodies can enhance repopulation of mature blood cell	
CC	linesages following chemotherapy, radiotherapy, or bone marrow	
CC	transplantation.	
XX		
SQ	Sequence 1000 AA;	
	Query Match 86.0%; Score 4533.5; DB 16; Length 1000;	
	Best Local Similarity 85.6%; Pred. No. 0;	
	Matches 851; Conservative 56; Mismatches 84; Indels 3; Gaps 2;	
QY	1 MPALA-RDAGTVLLVVFSAIFGTITNODLPVVKCVLHNKNDSSVYKSSYPMVSES 59	
DB	1 MRALAQRSDRLLLVLSVMILFTVNQDLPVKCVLISHENNGSSAGKPSYRMVRS 60	
QY	60 PEDLGCALRPQSSGTVYAAAVEVDVASITLQVLDVAPGNISCLWVFKHSLNCQPHFD 119	
DB	61 PEDLOCTPRQSEGTVEAATVEVAESGSILOVQATPGDLSCLWVFKHSLGCGPHFD 120	
QY	120 LQNRGVVSMILKMTQAGBYLLFIOSEATNYITLFTVSIIRNTLLYLRLRPYFKRMENQ 179	
DB	121 LQNRGIVSMALNVTQAGBYLLHIOSEAAANYTLFTVNVNVDQLYVLRPYFKRMENQ 180	
QY	180 DALVCISESVPEPIVENVLCDOSGESCKEESPAVYKKEKVLHFLGTDIRCCARNELGR 239	
DB	181 DALLCISEGVEPTVEWVLCSSHRESCKEGPAVYKKEKVLHFLGTDIRCCARNALGR 240	
QY	240 ECTRLEFTIDLNQTPQTTLPQLFLKVGPEPLWIRCKAVHNVHGFGLTWELKALEEGNYFE 299	
DB	241 ESTKLTIDLNQAQOSTLPQLFLKVGPEPLWIRCKAIHNVHGFGLTWELKALEEGSYFE 300	
QY	300 MSTYSTNRTMIRILFAFVSSVARNDTYTCSSSKHPSQSALVTIVGKGFINATNSSDY 359	
DB	301 MSTYSTNRTMIRILFAFVSSVGRNDTYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360	
QY	360 EIDQYEFPCFSVRFKAYPQIRCTWTFSRKFPCEQKGLDNGYSISKPCNHKHQPGEYIFH 419	
DB	361 EIDPYEKFCFSVRFKAYPRIRCTWTFISQASFPCEQKGLDNGYSISKPCDHKNKPGYIFY 420	
QY	420 AENDDAQFTKMFTLNIRKPOVLAAEASQASQCFSDGYPLPSWTWKKCSKSPNCTEIT 479	
DB	421 AENDDAQFTKMFTLNIRKPOVLAAEASQASQSDGYPLPSWTWKKCSKSPNCTEIP 480	
QY	480 EGVNRRKANRKFQGVSSSTLANMSEAIGFLVKCCAYNSLGTSCETILLNSPGPPFFIQ 539	
DB	481 EGVNRRKANRKFQGVSSSTLANMSEAIGFLVKCCAYNSMGTSCTIFLNSPGPPFFIQ 540	
QY	540 DNISFYATICVCLLFIIVLLTLLCHYKKOFRYESQLQMVQVGTSSDNEYFYVDFREY 599	
DB	541 DNISFYATICGLCLPFIIVLLTLLCHYKKOFRYESQLQMVQVGTSSDNEYFYVDFRDEY 600	
QY	600 DLKWEFPRENLEFGKVLGSGAFKVMNATAYISKTGVSIOQAVKMLKEKADSSEREA 659	
DB	601 DLKWEFPRENLEFGKVLGSGAFKVMNATAYISKTGVSIOQAVKMLKEKADSSEREA 660	
QY	660 SELKMMTQLGSHENIVNLLGACTLSGPIYLIFEYCCYGDLLNLRSKREKFRHTWTIEFK 719	
DB	661 SELKMMTHLGHHDNIVNLLGACTLSGPIYLIFEYCCYGDLLNLRSKREKFRHTWTIEFK 720	
QY	720 EHNFSYPTFQSHPNSSMPGSRVQTHPDSQISGLHGNFSHSEDETEYENQKRL--EEE 777	

Db	721 EHNFSYPTFOAHSSMPGSRVQLHPPLDQLSGFNGNSIHSDETEYENQKRLAESEE 780
QY	778 EDNLNLTEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIM 837
Db	781 EDNLNLTEDLLCFAYQVAKGMFELEFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
QY	838 SDSNVYVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 897
Db	841 SDSNVYVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDAN 900
QY	898 FYKLIQNGFKMDQPFYATAEIYIIQSWAFPSRRKPSFPNLTSLGCOLADAEAEAMYON 957
Db	901 FYKLIQSGFKMDQPFYATEGIIYFVMSQWAFDSRRKPSFPNLTSLGCOLADAEAEAMYON 960
QY	958 VDGRYSECPHTYQNRPRPFSREMDLGLLSPOAQVE 991
Db	961 MGGNYPEHPSIYQNRPRPLSREAGSEPPSPQAOVK 994

Search completed: May 27, 2003, 14:31:20  
Job time : 72.0307 secs

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:28:59 ; Search time 24.0121 Seconds  
(without alignments)  
1216.760 Million cell updates/sec

Title: US-09-919-408-4  
Perfect score: 5274  
Sequence: 1 MPALARDAGTVPLLVFSAM.....PFSREMDLGLLSPOAQVEDS 993

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTCUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5274	100.0	993	1	US-07-977-451-4
2	5274	100.0	993	1	US-08-252-517-4
3	5274	100.0	993	1	US-07-906-397A-4
4	5274	100.0	993	1	US-08-601-891-4
5	5274	100.0	993	2	US-09-021-324-4
6	5274	100.0	993	5	PCT-US92-09893-4
7	5274	100.0	1160	5	PCT-US92-03401-4
8	5240	99.4	993	1	US-08-222-299-4
9	5240	99.4	993	2	US-08-434-878-4
10	5240	99.4	993	5	PCT-US95-03718-4
11	5238	99.3	993	1	US-08-183-211-2
12	5238	99.3	993	5	PCT-US95-00176A-2
13	4333.5	86.0	1000	1	US-08-222-299-2
14	4333.5	86.0	1000	2	US-08-434-878-2
15	4333.5	86.0	1000	5	PCT-US95-03718-2
16	4429.5	84.0	992	1	US-07-813-593-2
17	4429.5	84.0	992	1	US-07-977-451-2
18	4429.5	84.0	992	1	US-07-946-507-2
19	4429.5	84.0	992	1	US-08-232-517-2
20	4429.5	84.0	992	1	US-07-906-397A-2
21	4429.5	84.0	992	1	US-08-601-891-2
22	4429.5	84.0	992	2	US-09-021-324-2
23	4429.5	84.0	992	5	PCT-US92-02750-2
24	4429.5	84.0	992	5	PCT-US92-03401-2
25	4429.5	84.0	992	5	PCT-US92-09893-2
26	1654	31.4	481	4	US-07-912-122-4
27	1654	31.4	481	5	PCT-US93-06404-4

28	1251	23.7	976	3	US-08-750-141A-1	Sequence 1, Appli
29	1216.5	23.1	972	3	US-08-750-141A-2	Sequence 2, Appli
30	1166	22.1	1089	1	US-08-180-195-36	Sequence 36, Appli
31	1166	22.1	1089	1	US-08-168-917-4	Sequence 4, Appli
32	1166	22.1	1089	1	US-08-477-329-36	Sequence 36, Appli
33	1166	22.1	1089	2	US-08-475-458-36	Sequence 36, Appli
34	1166	22.1	1089	2	US-08-460-510-4	Sequence 4, Appli
35	1166	22.1	1089	2	US-08-460-490-4	Sequence 4, Appli
36	1166	22.1	1089	3	US-08-980-400-36	Sequence 36, Appli
37	1166	22.1	1089	3	US-08-462-728-2	Sequence 2, Appli
38	1166	22.1	1089	4	US-09-583-458A-36	Sequence 36, Appli
39	1166	22.1	1089	4	US-09-583-210-36	Sequence 36, Appli
40	1166	22.1	1089	4	US-09-583-449A-36	Sequence 36, Appli
41	1166	22.1	1089	4	US-09-435-059-36	Sequence 36, Appli
42	1166	22.1	1089	4	US-08-461-917-2	Sequence 2, Appli
43	1166	22.1	1089	5	PCT-US92-00730-4	Sequence 4, Appli
44	1166	22.1	1089	5	PCT-US92-00862-4	Sequence 4, Appli
45	1086.5	20.6	1106	1	US-08-180-195-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-07-977-451-4  
; Sequence 4, Application US/07977451  
; Patent No. 5270458  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Imclone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/977,451  
; FILING DATE: 19921119  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US UNASSIGNED  
; FILING DATE: 12-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-977-451-4

Query Match 100.0%; Score 5274; DB 1; Length 993;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVFSAMIFGTITNODLPVKVCLINHNKNDSSVGKSSYPMVSESP 60  
DB 1 MPALARDAGTVPLLVFSAMIFGTITNODLPVKVCLINHNKNDSSVGKSSYPMVSESP 60  
QY 61 EDLCALRPQSSGTVYEAADVVSASITLOVLVDAPGNISCLVWFKHSSLNCPHFEDL 120  
DB 61 EDLCALRPQSSGTVYEAADVVSASITLOVLVDAPGNISCLVWFKHSSLNCPHFEDL 120  
QY 121 QNRGVSMVILKMTETQAGEVLLFIQSEATNYTILFTVSIKNTLLYLRPPYFRKMEQND 180  
DB 121 QNRGVSMVILKMTETQAGEVLLFIQSEATNYTILFTVSIKNTLLYLRPPYFRKMEQND 180  
QY 181 ALVCISEVPEPIVWVLCDSOGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
DB 181 ALVCISEVPEPIVWVLCDSOGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
QY 241 CTRLETDLNOTPQTLTLPOLFKVGEPLWIRCKAVVHNGFGLTWELNKALEEGNYFEM 300  
DB 241 CTRLETDLNOTPQTLTLPOLFKVGEPLWIRCKAVVHNGFGLTWELNKALEEGNYFEM 300  
QY 301 STYSTNRTMIRILFAFVSARNDGYTCSSKHPQSALVTIVGKGFINATNSSEYD 360  
DB 301 STYSTNRTMIRILFAFVSARNDGYTCSSKHPQSALVTIVGKGFINATNSSEYD 360  
QY 361 IDQYEEFCFSVRKAYPOIRCTWTFSRKFCEQKGLDNGYSISKFCNKHQGEYIPHA 420  
DB 361 IDQYEEFCFSVRKAYPOIRCTWTFSRKFCEQKGLDNGYSISKFCNKHQGEYIPHA 420  
QY 421 ENDDAQFTKMTLNRKRPQVLAESAQSCFSDGYPLPSWTWKCKSDKSPNCTEEITE 480  
DB 421 ENDDAQFTKMTLNRKRPQVLAESAQSCFSDGYPLPSWTWKCKSDKSPNCTEEITE 480  
QY 481 GWNKRNKRVFGQWVSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSGPPFFIOD 540  
DB 481 GWNKRNKRVFGQWVSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSGPPFFIOD 540  
QY 541 NISFYATIGVCLLFTLWTLTLLCHYKQFRYESQLOVMQVVTGSSDNEYFYDFREYED 600  
DB 541 NISFYATIGVCLLFTLWTLTLLCHYKQFRYESQLOVMQVVTGSSDNEYFYDFREYED 600  
QY 601 LKWEPPRENLEFGKVLGSAFGKVNATAYGISKTVGSIQVAVKMLKEKADSSEREALMS 660  
DB 601 LKWEPPRENLEFGKVLGSAFGKVNATAYGISKTVGSIQVAVKMLKEKADSSEREALMS 660  
QY 661 ELKMMTQLGSHENIYLLGACVLTSPVILFEYCCYGLLNLVRSKREKFRHTWTETPK 720  
DB 661 ELKMMTQLGSHENIYLLGACVLTSPVILFEYCCYGLLNLVRSKREKFRHTWTETPK 720  
QY 721 HNFSEYPTFQSHPNSSMPGSRVQIHPDSQISGLHGNFSHSEDETEYENQKRLSEEDL 780  
DB 721 HNFSEYPTFQSHPNSSMPGSRVQIHPDSQISGLHGNFSHSEDETEYENQKRLSEEDL 780

QY 781 NVLTFEDLLCFAYOVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840  
DB 781 NVLTFEDLLCFAYOVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840  
QY 841 NYVVRGNARLPVKWMAPESELFEIYTIKSDVMSYGILLWEIFSLGVNYPYGPVDPANFYK 900  
DB 841 NYVVRGNARLPVKWMAPESELFEIYTIKSDVMSYGILLWEIFSLGVNYPYGPVDPANFYK 900  
QY 901 LQNGFKMDQPFYATEEYIIMQSCWAFSDSKRPFNPNTLSLGCQLADAEAEANYQNVDG 960  
DB 901 LQNGFKMDQPFYATEEYIIMQSCWAFSDSKRPFNPNTLSLGCQLADAEAEANYQNVDG 960  
QY 961 RVSECPHTYQNRPRPESREMDLGLLSPOAQVEDS 993  
DB 961 RVSECPHTYQNRPRPESREMDLGLLSPOAQVEDS 993  
RESULT 2  
US-08-252-517-4  
; Sequence 4, Application US/08252517  
; Patent No. 5548065  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Imclone Systems Incorporated  
; STREET: 180 Varick Street  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10014  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/252,517  
; FILING DATE: 31-OCT-1994  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/977,451  
; FILING DATE: 19-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/906,397  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/05401  
; FILING DATE: 26-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: TW 81102961  
; FILING DATE: 15-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US PCT/US92/02750  
; FILING DATE: 02-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/813,593  
; FILING DATE: 24-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/793,065  
; FILING DATE: 15-NOV-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/728,913  
; FILING DATE: 28-JUN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/679,666  
; FILING DATE: 02-APR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feit, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: LEM-3-7P

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-252-517-4

Query Match 100.0%; Score 5274; DB 1; Length 993;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPPLLVESAMIFGTITNODLPVKCVLINHKNDSSVGKSSSPWSESP 60  
DB 1 MPALARDAGTVPPLLVESAMIFGTITNODLPVKCVLINHKNDSSVGKSSSPWSESP 60  
QY 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120  
DB 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120  
QY 121 QNRGVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSIINTLLYTLRRPYFRKMNQD 180  
DB 121 QNRGVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSIINTLLYTLRRPYFRKMNQD 180  
QY 181 ALVCISESPEPIVWVLCDSQGESCKEESPAVVYKKEVHLHFGYDIDRCCARNELGRE 240  
DB 181 ALVCISESPEPIVWVLCDSQGESCKEESPAVVYKKEVHLHFGYDIDRCCARNELGRE 240  
QY 241 CTRLEFDLNOTPQTLPQLFLKVGPELWIRKAVHVNHGFGLTWELENKALEBGNFEM 300  
DB 241 CTRLEFDLNOTPQTLPQLFLKVGPELWIRKAVHVNHGFGLTWELENKALEBGNFEM 300  
QY 301 STYSTNRTMIRILFAFVSSVARNDGTGYTSSSKHPQSQALVTIVGKGFNATNSSDEYE 360  
DB 301 STYSTNRTMIRILFAFVSSVARNDGTGYTSSSKHPQSQALVTIVGKGFNATNSSDEYE 360  
QY 361 IDOYEFCFSVRKAYPOIRCTWTFRSKSPCEQKGLDNGYSISKFNHRHQGEYIFHA 420  
DB 361 IDOYEFCFSVRKAYPOIRCTWTFRSKSPCEQKGLDNGYSISKFNHRHQGEYIFHA 420  
QY 421 ENDDAOFTRMFTLINRKPVLAEASASQSCFSDGYPPLSWTKKCDKSPNCTEITE 480  
DB 421 ENDDAOFTRMFTLINRKPVLAEASASQSCFSDGYPPLSWTKKCDKSPNCTEITE 480  
QY 481 GVNRRKANRVFCQWSSSTLNNSEAIKGLVKCCAYNSLGTSCETILLNSPGFPFIQD 540  
DB 481 GVNRRKANRVFCQWSSSTLNNSEAIKGLVKCCAYNSLGTSCETILLNSPGFPFIQD 540  
QY 541 NISFYATIGVCLLFIIVVLLTLLCHKKYKQPRYSQLOMVOVQTSSDNEYFYVDFREYED 600  
DB 541 NISFYATIGVCLLFIIVVLLTLLCHKKYKQPRYSQLOMVOVQTSSDNEYFYVDFREYED 600  
QY 601 LKWEFPRENLEFCGLGSAFGVMNATYIGISKTGYSIOVAVKMLKEKADSSERREALMS 660  
DB 601 LKWEFPRENLEFCGLGSAFGVMNATYIGISKTGYSIOVAVKMLKEKADSSERREALMS 660  
QY 661 ELKMMTQLGSHENIVNLLGACTLSGPIYILFECYCCYGDLLNLYLSRKREKFRHTWEIFKE 720  
DB 661 ELKMMTQLGSHENIVNLLGACTLSGPIYILFECYCCYGDLLNLYLSRKREKFRHTWEIFKE 720  
QY 721 HNFSPYPTQSHNSMPGSRREVQIHPDSQISGLHGNFSHSEDETEYENQKRLSEEDL 780  
DB 721 HNFSPYPTQSHNSMPGSRREVQIHPDSQISGLHGNFSHSEDETEYENQKRLSEEDL 780  
QY 781 NVLTFFEDLLCFAYOVAKGMEFLFKSCVHRDLAARNVLTGHGVKVKICDFGLARDIMS 840  
DB 781 NVLTFFEDLLCFAYOVAKGMEFLFKSCVHRDLAARNVLTGHGVKVKICDFGLARDIMS 840  
QY 841 NYVVRGNARLPVKWMAPESLFEGITYIKSDVWSYGILLWEIFSLGPNYPGIPVDANFYK 900

DB 841 NYVVRGNARLPVKWMAPESLFEGITYIKSDVWSYGILLWEIFSLGPNYPGIPVDANFYK 900  
QY 901 LIQNGFKMDOPFYATEEYIIIMQSCWAFDSKRSPFNLTSLFLGCOLADAEAMYNVDG 960  
DB 901 LIQNGFKMDOPFYATEEYIIIMQSCWAFDSKRSPFNLTSLFLGCOLADAEAMYNVDG 960  
QY 961 RVSECPHTYQNRPFPSREMDLGLLSPOAQVEDS 993  
DB 961 RVSECPHTYQNRPFPSREMDLGLLSPOAQVEDS 993

RESULT 3  
US-07-906-397A-4  
Sequence 4, Application US/07906397A  
Patent No. 5621090  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCLONE SYSTEMS INCORPORATED  
STREET: 180 VARICK STREET  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/906,397A  
FILING DATE: 19920626  
CLASSIFICATION: 536  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-PPPPPP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-906-397A-4

Query Match 100.0%; Score 5274; DB 1; Length 993;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPPLLVESAMIFGTITNODLPVKCVLINHKNDSSVGKSSSPWSESP 60  
DB 1 MPALARDAGTVPPLLVESAMIFGTITNODLPVKCVLINHKNDSSVGKSSSPWSESP 60  
QY 61 EDLGCALRPOSSGTVYEAADVEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120

61 EDLGCALRPQSSGTVEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120  
121 QNRGVVSMVILKMTQAGEYLLFQSEATNTILFTVSIINTLLYLRPPYFRKMENOD 180  
121 QNRGVVSMVILKMTQAGEYLLFQSEATNTILFTVSIINTLLYLRPPYFRKMENOD 180  
181 ALVCISESVPPIVWLVKCDGSGESCKEESPAVVKKEKVLHFLFCTDIRCCARNELGRE 240  
181 ALVCISESVPPIVWLVKCDGSGESCKEESPAVVKKEKVLHFLFCTDIRCCARNELGRE 240  
241 CTRFLTIDLNQPTTLPQLFKVGEPLWIRCKAVHVNHGFGLTWELNKALEEGNYFEM 300  
241 CTRFLTIDLNQPTTLPQLFKVGEPLWIRCKAVHVNHGFGLTWELNKALEEGNYFEM 300  
301 STYSTNRTMIRILFAFVSSVARNDTGYTSCSSKHPQSALVTIVGKGFNATNSSDEYE 360  
301 STYSTNRTMIRILFAFVSSVARNDTGYTSCSSKHPQSALVTIVGKGFNATNSSDEYE 360  
361 IDOYEFCFSVRFKAYPQIRCTWTFRSKFPCEQKGLDNGYSISKFCNKHQPGYIFHA 420  
361 IDOYEFCFSVRFKAYPQIRCTWTFRSKFPCEQKGLDNGYSISKFCNKHQPGYIFHA 420  
421 ENDDAQTFTKMTLNIRRKQVLAESAASQSCFSDGYPLPSWTWKKCSKSPNCTEITE 480  
421 ENDDAQTFTKMTLNIRRKQVLAESAASQSCFSDGYPLPSWTWKKCSKSPNCTEITE 480  
481 GVNRRKANRVQGVVSSSTLNNSBAIKGLVKCCAYNSLGTSCETILLNSPGFPFIQD 540  
481 GVNRRKANRVQGVVSSSTLNNSBAIKGLVKCCAYNSLGTSCETILLNSPGFPFIQD 540  
541 NISFYATIGVCLLFIVVLTLLCHKKYKQFPYESQVMQVQVTSQSSONEFYVDFREYED 600  
541 NISFYATIGVCLLFIVVLTLLCHKKYKQFPYESQVMQVQVTSQSSONEFYVDFREYED 600  
601 LKWEFPRENLEFGKVLGSGAFGVMNATAYGISTGVSIOVAVKMLKEKADSSEREALMS 660  
601 LKWEFPRENLEFGKVLGSGAFGVMNATAYGISTGVSIOVAVKMLKEKADSSEREALMS 660  
661 ELKMTQLGSHENIVNLLGACTLSGPIYLLIFECYCGDNLNLYLSRREKPHRTWTEIFKE 720  
661 ELKMTQLGSHENIVNLLGACTLSGPIYLLIFECYCGDNLNLYLSRREKPHRTWTEIFKE 720  
721 HNFSEFTFQSHNSNMPGSGREVIQHPDSQISGLHNSFHSDEIEYENQKLEEBEDL 780  
721 HNFSEFTFQSHNSNMPGSGREVIQHPDSQISGLHNSFHSDEIEYENQKLEEBEDL 780  
781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVVYKICDFGLARDIMSDS 840  
781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVVYKICDFGLARDIMSDS 840  
841 NYVVRGNARLPVKWMAPELSEFEGYITKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900  
841 NYVVRGNARLPVKWMAPELSEFEGYITKSDVWSYGILLWEIFSLGVNYPGIPVDANFYK 900  
901 LIQNGFKMDQPFATBEIYIIMQSCWAFDGRKPSFNLTSFLGCQLADAEAEAMYQNVGD 960  
901 LIQNGFKMDQPFATBEIYIIMQSCWAFDGRKPSFNLTSFLGCQLADAEAEAMYQNVGD 960  
961 RVSECPHTYQNRFRPSREMDLGLLSPQAQVEDS 993  
961 RVSECPHTYQNRFRPSREMDLGLLSPQAQVEDS 993

RESULT 4

US-08-601-891-4  
: Sequence 4, Application US/08601891  
: Patent No. 5747651  
: GENERAL INFORMATION:  
: APPLICANT: Lemischka, Ihor R.  
: TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
: TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
: NUMBER OF SEQUENCES: 10  
: CORRESPONDENCE ADDRESS:

ADDRESSEE: ImClone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/601,891  
FILING DATE: 15-FEB-1996  
CLASSIFICATION: 530  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/977,451  
FILING DATE: 19-NOV-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 28-JUN-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Feit, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-601-891-4

Query Match 100.0%; Score 5274; DB 1; Length 993;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVWFSAIMFGTITNQDLPVIRKVLINHKNDSSVKGSSYPMVSESP 60  
DB 1 MPALARDAGTVPLLVWFSAIMFGTITNQDLPVIRKVLINHKNDSSVKGSSYPMVSESP 60  
QY 61 EDLGCALRPQSSGTVEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120  
DB 61 EDLGCALRPQSSGTVEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSSLNCPHFDL 120  
QY 121 QNRGVVSMVILKMTQAGEYLLFQSEATNTILFTVSIINTLLYLRPPYFRKMENOD 180  
DB 121 QNRGVVSMVILKMTQAGEYLLFQSEATNTILFTVSIINTLLYLRPPYFRKMENOD 180



```
QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240
Db 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240
QY 241 CTRLFTIDLNOTPOTTLPOLFLKVGEPDLWIRCKAVHNHGFGLTWELNKALEEGNFEM 300
Db 241 CTRLFTIDLNOTPOTTLPOLFLKVGEPDLWIRCKAVHNHGFGLTWELNKALEEGNFEM 300
QY 301 STYSTNTRMIRILFAFVSSVARNDTGYTSCSSKHPQSOSALVTIVGKGFINATNSSEDEYE 360
Db 301 STYSTNTRMIRILFAFVSSVARNDTGYTSCSSKHPQSOSALVTIVGKGFINATNSSEDEYE 360
QY 361 IDOYEFCFSVREKAYPOIRCTWTFRSKSPCEOKGLDNGYSISKFCNKHKHPGCEYIFHA 420
Db 361 IDOYEFCFSVREKAYPOIRCTWTFRSKSPCEOKGLDNGYSISKFCNKHKHPGCEYIFHA 420
QY 421 ENDDAQTFTKFTLNIRKPKQVLAESAQSCFSDGYPPLSWTWKKGSDKSPNCTEITE 480
Db 421 ENDDAQTFTKFTLNIRKPKQVLAESAQSCFSDGYPPLSWTWKKGSDKSPNCTEITE 480
QY 481 GVNRRKANRVFGQWSSSTLNKSEAIGFLVKCCAYNSLGTSCETILLNSPGPFPIOD 540
Db 481 GVNRRKANRVFGQWSSSTLNKSEAIGFLVKCCAYNSLGTSCETILLNSPGPFPIOD 540
QY 541 NISFYATIGVCLLFIVVLTLLCHIKYKQFYESOLQWQVGTSSDNEYFYVDFREYED 600
Db 541 NISFYATIGVCLLFIVVLTLLCHIKYKQFYESOLQWQVGTSSDNEYFYVDFREYED 600
QY 601 LKWEFFRENLEFGKVLGSGAFGKVMNATAYGISTGTYSIQVAVKMLKEKADSSREALMS 660
Db 601 LKWEFFRENLEFGKVLGSGAFGKVMNATAYGISTGTYSIQVAVKMLKEKADSSREALMS 660
QY 661 ELKMTQOLGSHENIVNLGACTLSGPYILIFEXCCYGDLLNLYLSRKREKPHRTWTEIFKE 720
Db 661 ELKMTQOLGSHENIVNLGACTLSGPYILIFEXCCYGDLLNLYLSRKREKPHRTWTEIFKE 720
QY 721 HNFSEFYTFQSHPNSSMPGSRVQIHPDSQISGLHNSFHSDEIEYENOKRLEEBEDL 780
Db 721 HNFSEFYTFQSHPNSSMPGSRVQIHPDSQISGLHNSFHSDEIEYENOKRLEEBEDL 780
QY 781 NVLTFEDLLCFAYQVAKGMFLEFKSCVHRDLAARNVLVTHGKVVKTCDFGLARDIMSDS 840
Db 781 NVLTFEDLLCFAYQVAKGMFLEFKSCVHRDLAARNVLVTHGKVVKTCDFGLARDIMSDS 840
QY 841 NYVVRGNARLPVKWMAPESLFEGYITIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
Db 841 NYVVRGNARLPVKWMAPESLFEGYITIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
QY 901 LIQNGFKMDOPFYATEEIIYIMOSWAFDSRKRPSFPNLTSLGCLQADAEAEAMYQNVDG 960
Db 901 LIQNGFKMDOPFYATEEIIYIMOSWAFDSRKRPSFPNLTSLGCLQADAEAEAMYQNVDG 960
QY 961 RVSECPHTYONRRPFRREMDLGLLSPOAQVEDS 993
Db 961 RVSECPHTYONRRPFRREMDLGLLSPOAQVEDS 993

RESULT 5
US-09-021-324-4
; Sequence 4, Application US/09021324
; Patent No. 5912133
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSES: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/021,324
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 1992-11-19
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 993 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-021-324-4
```

```
Query Match 100.0%; Score 5274; DB 2; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MPALARDAGTVPDLLVVFESAMIFGTITNODLPVVKVLINHNKNDSSVCKSSSYPMWSESP 60
Db 1 MPALARDAGTVPDLLVVFESAMIFGTITNODLPVVKVLINHNKNDSSVCKSSSYPMWSESP 60
QY 61 EDLGCALRPOSSGTVYEAATAVEVDVSASITLQVLVDAPGNISCLWFKHSSLNCPHFDL 120
Db 61 EDLGCALRPOSSGTVYEAATAVEVDVSASITLQVLVDAPGNISCLWFKHSSLNCPHFDL 120
QY 121 QNRGVWSMVLKMTETQAGEYLLFIQSEATNYTLFTVSIRNTLLYTLRRPYFRKMNQD 180
Db 121 QNRGVWSMVLKMTETQAGEYLLFIQSEATNYTLFTVSIRNTLLYTLRRPYFRKMNQD 180
QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240
Db 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240
QY 241 CTRLFTIDLNOTPOTTLPOLFLKVGEPDLWIRCKAVHNHGFGLTWELNKALEEGNFEM 300
Db 241 CTRLFTIDLNOTPOTTLPOLFLKVGEPDLWIRCKAVHNHGFGLTWELNKALEEGNFEM 300
```

```

Db 241 CTRLFTIDLQTPQTTLPOLFLKVGEPPLIRKAVHVNHGFLTWELNKALEEGNYFEM 300
Qy 301 STYSTNRTMIRILFAFVSSVARNDTYGTCSSSKHPSQSALVTIVGKGFNATNSSEDEY 360
Db 301 STYSTNRTMIRILFAFVSSVARNDTYGTCSSSKHPSQSALVTIVGKGFNATNSSEDEY 360
Qy 361 IDQYEEFCFSVRKAYPQIRCTWTFSRKSPPCQKGLDNGYSISKFCNKHQPGYIFHA 420
Db 361 IDQYEEFCFSVRKAYPQIRCTWTFSRKSPPCQKGLDNGYSISKFCNKHQPGYIFHA 420
Qy 421 ENDDAQFTKMTLNIIRKPOVLAESAASQASCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
Db 421 ENDDAQFTKMTLNIIRKPOVLAESAASQASCFSDGYPLPSWTWKCKSDKSPNCTEITE 480
Qy 481 GWNKRKANRKFVGQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
Db 481 GWNKRKANRKFVGQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPFPIQD 540
Qy 541 NISFYATIGVCLLFIVVLTLLCHKYYKKQFYESQOLQWQVGTGSSDNEYFYVDFREYED 600
Db 541 NISFYATIGVCLLFIVVLTLLCHKYYKKQFYESQOLQWQVGTGSSDNEYFYVDFREYED 600
Qy 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIOVAVKMLKEKADSSEREALMS 660
Db 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIOVAVKMLKEKADSSEREALMS 660
Qy 661 ELKMMTQLGSHENIVNLLGACTLSGPIYILIFEYCCYDGLLNLNLSRKREKPHRTWTEIFE 720
Db 661 ELKMMTQLGSHENIVNLLGACTLSGPIYILIFEYCCYDGLLNLNLSRKREKPHRTWTEIFE 720
Qy 721 HNFsfyPTFQSHPNSSMPGSRVQIHPDSQOISGLHGNFSFSEDEIYENOKRLEEDL 780
Db 721 HNFsfyPTFQSHPNSSMPGSRVQIHPDSQOISGLHGNFSFSEDEIYENOKRLEEDL 780
Qy 781 NVLTFEDLLCFAYQAVAKMEFLFKSCVHRDLAARNVLYTHGKVVYKICDFGLARDIMSDS 840
Db 781 NVLTFEDLLCFAYQAVAKMEFLFKSCVHRDLAARNVLYTHGKVVYKICDFGLARDIMSDS 840
Qy 841 NYVVRGNARLPVKWMAPESLFEGIIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
Db 841 NYVVRGNARLPVKWMAPESLFEGIIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900
Qy 901 LQIONGFKMQDPFYATBEIYIIMOSWAFDSKRKPSFNLTSFLGCCQADAEAMYNQVNDG 960
Db 901 LQIONGFKMQDPFYATBEIYIIMOSWAFDSKRKPSFNLTSFLGCCQADAEAMYNQVNDG 960
Qy 961 RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS 993
Db 961 RVSECPHTYQNRPPFSREMDLGLLSPOAQVEDS 993

```

## RESULT 6

PCT-US92-09893-4

; Sequence 4, Application PC/TUS9209893

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ibor R.

; TITLE OF INVENTION: TOPIPOTENT HEMATOPOIETIC STEM CELL

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ImClone Systems Incorporated

; STREET: 180 Varick Street

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10014

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/09893

; FILING DATE: 19921116

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Feit, Irving N.

; REGISTRATION NUMBER: 28,601

; REFERENCE/DOCKET NUMBER: LEM-3-7PT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-645-1405

; TELEFAX: 212-645-2054

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 993 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; PCT-US92-09893-4

Query Match 100.0%; Score 5274; DB 5; Length 993;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPALARDAGTVPLLVVPSAMIFGTITNODLPVVKVLINHKNDSSVKGSSSPMVWSESP 60

Db 1 MPALARDAGTVPLLVVPSAMIFGTITNODLPVVKVLINHKNDSSVKGSSSPMVWSESP 60

Qy 61 EDLGCALRPOSSGTVEAAAVEVDVSASITLOVLVDAPGNISCLWVKFHSLSLNCQPHFDL 120

Db 61 EDLGCALRPOSSGTVEAAAVEVDVSASITLOVLVDAPGNISCLWVKFHSLSLNCQPHFDL 120

Qy 121 QNRGVWSMWILKMTQAGEYLLFIQSEATNYTLFTVSIIRNTLLYTLRRPYFRKMENQD 180

Db 121 QNRGVWSMWILKMTQAGEYLLFIQSEATNYTLFTVSIIRNTLLYTLRRPYFRKMENQD 180

Qy 181 ALVCISSEVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240

Db 181 ALVCISSEVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240

Qy 241 CTRLFTIDLQTPQTTLPOLFLKVGEPPLIRKAVHVNHGFLTWELNKALEEGNYFEM 300

Db 241 CTRLFTIDLQTPQTTLPOLFLKVGEPPLIRKAVHVNHGFLTWELNKALEEGNYFEM 300

Qy 301 STYSTNRTMIRILFAFVSSVARNDTYGTCSSSKHPSQSALVTIVGKGFNATNSSEDEY 360

Db 301 STYSTNRTMIRILFAFVSSVARNDTYGTCSSSKHPSQSALVTIVGKGFNATNSSEDEY 360

Qy 361 IDQYEEFCFSVRKAYPQIRCTWTFSRKSPPCQKGLDNGYSISKFCNKHQPGYIFHA 420

Db 361 IDQYEEFCFSVRKAYPQIRCTWTFSRKSPPCQKGLDNGYSISKFCNKHQPGYIFHA 420

Qy 421 ENDDAQFTKMTLNIIRKPOVLAESAASQASCFSDGYPLPSWTWKCKSDKSPNCTEITE 480

Db 421 ENDDAQFTKMTLNIIRKPOVLAESAASQASCFSDGYPLPSWTWKCKSDKSPNCTEITE 480

Qy 481 GWNKRKANRKFVGQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPFPIQD 540

Db 481 GWNKRKANRKFVGQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPFPIQD 540

Qy 541 NISFYATIGVCLLFIVVLTLLCHKYYKKQFYESQOLQWQVGTGSSDNEYFYVDFREYED 600

Db 541 NISFYATIGVCLLFIVVLTLLCHKYYKKQFYESQOLQWQVGTGSSDNEYFYVDFREYED 600

Qy 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIOVAVKMLKEKADSSEREALMS 660

Db 601 LKWEFPRENLEFGKVLGSGAFGKVMNATAYGISKTVGSIOVAVKMLKEKADSSEREALMS 660

Qy 661 ELKMMTQLGSHENIVNLLGACTLSGPIYILIFEYCCYDGLLNLNLSRKREKPHRTWTEIFE 720

Db 661 ELKMMTQLGSHENIVNLLGACTLSGPIYILIFEYCCYDGLLNLNLSRKREKPHRTWTEIFE 720

Qy 721 HNFsfyPTFQSHPNSSMPGSRVQIHPDSQOISGLHGNFSFSEDEIYENOKRLEEDL 780

Db 721 HNFsfyPTFQSHPNSSMPGSRVQIHPDSQOISGLHGNFSFSEDEIYENOKRLEEDL 780

QY 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVVKICDFGLARDIMSDS 840  
DB 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVVKICDFGLARDIMSDS 840  
QY 841 NYVVGNGARLPVKWMAPELFEIYITKSDVMSYGILLWEIFSLGVNPGIPVDANFYK 900  
DB 841 NYVVGNGARLPVKWMAPELFEIYITKSDVMSYGILLWEIFSLGVNPGIPVDANFYK 900  
QY 901 LIQNGFKMDQPFYATEEYIIIMQSCWAFDSKRPSFPNLTSLFGLCOLADAEAMYQNVDG 960  
DB 901 LIQNGFKMDQPFYATEEYIIIMQSCWAFDSKRPSFPNLTSLFGLCOLADAEAMYQNVDG 960  
QY 961 RVSECPHTYQNRPRPSREMDLGLLSPQAQVEDS 993  
DB 961 RVSECPHTYQNRPRPSREMDLGLLSPQAQVEDS 993

## RESULT 7

PCR-US92-05401-4  
; Sequence 4, Application PC/TUS9205401  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: IMCLONE SYSTEMS INCORPORATED  
; STREET: 180 VARICK STREET  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: U.S.A.  
; ZIP: 10014

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/05401  
; FILING DATE: 19920626  
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
; NAME: Felt, Irving N.  
; REGISTRATION NUMBER: 28,601  
; REFERENCE/DOCKET NUMBER: LEM-3-PPPPPT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-645-1405  
; TELEFAX: 212-645-2054

INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1160 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US92-05401-4

Query Match 100.0%; Score 5274; DB 5; Length 1160;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVVFSAMIFGTITNQDLPVVKVCLINHKNDSSVKGKSSYPMVSESP 60  
DB 20 MPALARDAGTVPLLVVFSAMIFGTITNQDLPVVKVCLINHKNDSSVKGKSSYPMVSESP 79  
QY 61 EDLGCALRQSSGTVYEAAREVDVDSASITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120  
DB 80 EDLGCALRQSSGTVYEAAREVDVDSASITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 139  
QY 121 QNRGVVSMVLKMTETQAGEYLLFIQSEATNYTILFTVSRNTLLYTLRPPYFKMENQD 180  
DB 140 QNRGVVSMVLKMTETQAGEYLLFIQSEATNYTILFTVSRNTLLYTLRPPYFKMENQD 199  
QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 240

DB 200 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFGTDIRCCARNELGRE 259  
QY 241 CTRUFTIDLNOTPQTTLPQLFLKVGEPLEWIRCKAVVHVHNGFGLTWELNKALEEGNPFM 300  
DB 260 CTRUFTIDLNOTPQTTLPQLFLKVGEPLEWIRCKAVVHVHNGFGLTWELNKALEEGNPFM 319  
QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSOSALVTIVGKGFINATNSEDYE 360  
DB 320 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSOSALVTIVGKGFINATNSEDYE 379  
QY 361 IDQYEEFCFSVRKAYPOIRCTWTFSRKSPFCEQKGLDNGYSISKFCNKKHQPGEYIFHA 420  
DB 380 IDQYEEFCFSVRKAYPOIRCTWTFSRKSPFCEQKGLDNGYSISKFCNKKHQPGEYIFHA 439  
QY 421 ENDDAQFTKMETLNIIRKPOVLAEASASQSCFSQGYPLPSPWTKKCDKSPNCTEETE 480  
DB 440 ENDDAQFTKMETLNIIRKPOVLAEASASQSCFSQGYPLPSPWTKKCDKSPNCTEETE 499  
QY 481 GWNRKANKRVFGOWVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGEFFIQD 540  
DB 500 GWNRKANKRVFGOWVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGEFFIQD 559  
QY 541 NISFYATIGVCLLFIVLTLTLLCHIKYKQFRYESQLOVQVOTGSSDNEYFYVDFREYED 600  
DB 560 NISFYATIGVCLLFIVLTLTLLCHIKYKQFRYESQLOVQVOTGSSDNEYFYVDFREYED 619  
QY 601 LKWEFPRENLEFGKVGSGAFGKVMNATAYGISTGVSIQVAVKMLKEKADSSREALMS 660  
DB 620 LKWEFPRENLEFGKVGSGAFGKVMNATAYGISTGVSIQVAVKMLKEKADSSREALMS 679  
QY 661 ELKMMTOLGSHENIVNLLGACTLSGPIYLIFFEYCCYGOLLNLYLSKRKFKHRTWTEIFKE 720  
DB 680 ELKMMTOLGSHENIVNLLGACTLSGPIYLIFFEYCCYGOLLNLYLSKRKFKHRTWTEIFKE 739  
QY 721 HNF3FYPTFQSHPNSSMPGSRVQIHDPDSDOIISGLHGNFSHSEDEIEYENOKRLEEDL 780  
DB 740 HNF3FYPTFQSHPNSSMPGSRVQIHDPDSDOIISGLHGNFSHSEDEIEYENOKRLEEDL 799  
QY 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVVKICDFGLARDIMSDS 840  
DB 800 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVVKICDFGLARDIMSDS 859  
QY 841 NYVVGNGARLPVKWMAPELFEIYITKSDVMSYGILLWEIFSLGVNPGIPVDANFYK 900  
DB 860 NYVVGNGARLPVKWMAPELFEIYITKSDVMSYGILLWEIFSLGVNPGIPVDANFYK 919  
QY 901 LIQNGFKMDQPFYATEEYIIIMQSCWAFDSKRPSFPNLTSLFGLCOLADAEAMYQNVDG 960  
DB 920 LIQNGFKMDQPFYATEEYIIIMQSCWAFDSKRPSFPNLTSLFGLCOLADAEAMYQNVDG 979  
QY 961 RVSECPHTYQNRPRPSREMDLGLLSPQAQVEDS 993  
DB 980 RVSECPHTYQNRPRPSREMDLGLLSPQAQVEDS 1012

## RESULT 8

US-08-222-299-4  
; Sequence 4, Application US/08222299  
; Patent No. 5635388  
; GENERAL INFORMATION:

APPLICANT: Bennett, Brian D.  
APPLICANT: Broz, Susan D.  
APPLICANT: Matthews, William  
APPLICANT: Zeigler, Francis C.  
TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA

ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/222,299  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 879  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-222-299-4

Query Match 99.4%; Score 5240; DB 1; Length 993;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 987; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
QY 1 MPALARDAGTVPVLLVFSAMIFGTITNODLPVVKVCLINHKNDSSVGVKSSYPMVSESP 60  
DB 1 MPALARDGGQLPLLVFSAMIFGTITNODLPVVKVCLINHKNDSSVGVKSSYPMVSESP 60  
QY 61 EDLGCALRPOSSGTVYEAADVDSASITLOVLVDAPGNISCLWFKHSLNCQPHFDL 120  
DB 61 EDLGCALRPOSSGTVYEAADVDSASITLOVLVDAPGNISCLWFKHSLNCQPHFDL 120  
QY 121 QNRGVSMVILKMTQAGEYLLFIQSEATNYTLFTVVSIRNTLLYTLRRPYFRKMENQD 180  
DB 121 QNRGVSMVILKMTQAGEYLLFIQSEATNYTLFTVVSIRNTLLYTLRRPYFRKMENQD 180  
QY 181 ALVCISSEVPPIVEWVLCDSQSGESKESPAVVKKEKVLHFLFGTDIRCARNELGRE 240  
DB 181 ALVCISSEVPPIVEWVLCDSQSGESKESPAVVKKEKVLHFLFGTDIRCARNELGRE 240  
QY 241 CTRLTIDLNOTPTTLPLQFLKVGEPILWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300  
DB 241 CTRLTIDLNOTPTTLPLQFLKVGEPILWIRCKAVHVNHGFLTWELNKALEEGNYFEM 300  
QY 301 STYSTNTRMIRLFAFVSSVARNDTGYTSCSSKHPSQSALVTIVGKGFINATNSSEDE 360  
DB 301 STYSTNTRMIRLFAFVSSVARNDTGYTSCSSKHPSQSALVTIVGKGFINATNSSEDE 360  
QY 361 IDQYEEFCFVRKAYQIRCTWTFSRKSFCQKGLDNGYSISKFCNKHOPGEYIFHA 420  
DB 361 IDQYEEFCFVRKAYQIRCTWTFSRKSFCQKGLDNGYSISKFCNKHOPGEYIFHA 420  
QY 421 ENDDAQFTKMTLIRRRKPVQLAASASOACFSFGYPLPSWTWKCSKSPNCTEEITE 480  
DB 421 ENDDAQFTKMTLIRRRKPVQLAASASOACFSFGYPLPSWTWKCSKSPNCTEEITE 480  
QY 481 GWNRRKANRVFGQWVSSSTLNMSAIGFLVKCAYNSLTGTCETILLNSGPPFPFQD 540  
DB 481 GWNRRKANRVFGQWVSSSTLNMSAIGFLVKCAYNSLTGTCETILLNSGPPFPFQD 540  
QY 541 NISFYATIGVCLLFIIVLTLLICHYKKQFRYESOLOMVQVGTSSDNEYFYVDFREYED 600  
DB 541 NISFYATIGVCLLFIIVLTLLICHYKKQFRYESOLOMVQVGTSSDNEYFYVDFREYED 600  
QY 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISKTGVSIOQVAVKMLKEKADSSREALMS 660

DB 601 LKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISKTGVSIOQVAVKMLKEKADSSREALMS 660  
QY 661 ELKMMTQLGSHENIVNLLGACTLSGGPIYLIFEXCCYGDLLNLYRSKRKFHRTWTEIFKE 720  
DB 661 ELKMMTQLGSHENIVNLLGACTLSGGPIYLIFEXCCYGDLLNLYRSKRKFHRTWTEIFKE 720  
QY 721 HNFSEFPTFQSHPNSSMPGSRVQIHPDSDQISGLHNSFHSDEIEYENQKRLEEEEDL 780  
DB 721 HNFSEFPTFQSHPNSSMPGSRVQIHPDSDQISGLHNSFHSDEIEYENQKRLEEEEDL 780  
QY 781 NVLTFFEDLLCFAYQVAKGMEFLFEKSCVHRDLAARNVLVTHGKVVKICDFGLARDINSDS 840  
DB 781 NVLTFFEDLLCFAYQVAKGMEFLFEKSCVHRDLAARNVLVTHGKVVKICDFGLARDINSDS 840  
QY 841 NYVVRGNARLPVKWMAPELSLFEGIYITKSDVMSYGILLWEIFSLGVNVPYGPPIVDANFYK 900  
DB 841 NYVVRGNARLPVKWMAPELSLFEGIYITKSDVMSYGILLWEIFSLGVNVPYGPPIVDANFYK 900  
QY 901 LIQNGFKMDQPPYATEEIIYIMQSCWAFDSKRKPSFPNLTSLFGLCQLADAEAEAMYQNVDG 960  
DB 901 LIQNGFKMDQPPYATEEIIYIMQSCWAFDSKRKPSFPNLTSLFGLCQLADAEAEAMYQNVDG 960  
QY 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993  
DB 961 RVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993

RESULT 9  
US-08-434-878-4  
Sequence 4, Application US/08434878  
Patent No. 5997865  
GENERAL INFORMATION:  
APPLICANT: Bennett, Brian D.  
APPLICANT: Broz, Susan D.  
APPLICANT: Matthews, William  
APPLICANT: Zeigler, Francis C.  
TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/434,878  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 879  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-434-878-4

Query Match 99.4%; Score 5240; DB 2; Length 993;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 987; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPVLLVFSAMIFGTITNODLPVVKVCLINHNKNDSSVGKSSSYPMVSESP 60  
DB 1 MPALARDGGQLPVLVFSAMIFGTITNODLPVVKVCLINHNKNDSSVGKSSSYPMVSESP 60

QY 61 EDLGCALRPOSSGTVYEAADVDSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120  
DB 61 EDLGCALRPOSSGTVYEAADVDSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120

QY 121 QNRGVVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSIINTLLYTLRRPYFRKMENQD 180  
DB 121 QNRGVVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSIINTLLYTLRRPYFRKMENQD 180

QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240  
DB 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240

QY 241 CTRLFTIDLNOTPQTTLPQLFKVGEPLWIRCKAVHVNHGFGLTWELNKALEEGNYFEM 300  
DB 241 CTRLFTIDLNOTPQTTLPQLFKVGEPLWIRCKAVHVNHGFGLTWELNKALEEGNYFEM 300

QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTTCSSSKHPQSOSALVTIVGKGFINATNSSEDEYE 360  
DB 301 STYSTNRTMIRILFAFVSSVARNDTGYTTCSSSKHPQSOSALVTIVGKGFINATNSSEDEYE 360

QY 361 IDOYEEFCFSVRKAYPOIRCTWTFSRKSPCEOKGLDNGYSISKFCNKHKQHPGEYIFHA 420  
DB 361 IDOYEEFCFSVRKAYPOIRCTWTFSRKSPCEOKGLDNGYSISKFCNKHKQHPGEYIFHA 420

QY 421 ENDDAOFKMTLNIRKPOVLAESAASQSCFSDGYPSPSWTKKSDKSPNCTEITE 480  
DB 421 ENDDAOFKMTLNIRKPOVLAESAASQSCFSDGYPSPSWTKKSDKSPNCTEITE 480

QY 481 GVNRRKANRKFQWVSSSTLNKSEALKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540  
DB 481 GVNRRKANRKFQWVSSSTLNKSEALKGLVKCCAYNSLGTSCETILLNSPGPFPIQD 540

QY 541 NISFYATIGVCLLFIIVVLLILCHYKQPRYESOLOMQVQVTGSSDNEYFYDFREYED 600  
DB 541 NISFYATIGVCLLFIIVVLLILCHYKQPRYESOLOMQVQVTGSSDNEYFYDFREYED 600

QY 601 LKWEFPRENLEFGKVLGSAFGKVMNATAGISKTGVSIOVAVKMLKEKADSSERREALMS 660  
DB 601 LKWEFPRENLEFGKVLGSAFGKVMNATAGISKTGVSIOVAVKMLKEKADSSERREALMS 660

QY 661 ELKMMTQLGSHENIVNLGACTLSGPYLLIFCYCCYDGLNLYLSRKEKPHRTWTEIFKE 720  
DB 661 ELKMMTQLGSHENIVNLGACTLSGPYLLIFCYCCYDGLNLYLSRKEKPHRTWTEIFKE 720

QY 721 HNFSPYPTQSHPNMSPGSRVQIHPDSQISGLHGNPSHSEDEIYEYENOKRLEEBEDL 780  
DB 721 HNFSPYPTQSHPNMSPGSRVQIHPDSQISGLHGNPSHSEDEIYEYENOKRLEEBEDL 780

QY 781 NVLTFEDLLCFAQVAKGMEFLFKSCVHRDLAARNVLYTHGKVVKICDPLGLARDIMSDS 840  
DB 781 NVLTFEDLLCFAQVAKGMEFLFKSCVHRDLAARNVLYTHGKVVKICDPLGLARDIMSDS 840

QY 841 NVYVRGNARLPVKWMAPESLFEGIYITKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900  
DB 841 NVYVRGNARLPVKWMAPESLFEGIYITKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900

QY 901 LIQNGFKMDOPFYATEEIIYIMOSWAFDSRKRSPNLTSLFGCOLADAEAMQNVGD 960  
DB 901 LIQNGFKMDOPFYATEEIIYIMOSWAFDSRKRSPNLTSLFGCOLADAEAMQNVGD 960

QY 961 RVSECPHTTQNRPPFSREMDLGLLSPQAQVEDS 993  
DB 961 PVSECPHTYQNRPPFSREMDLGLLSPQAQVEDS 993

RESULT 10  
PCT-US95-03718-4  
Sequence 4, Application PC/TUS9503718  
GENERAL INFORMATION:  
APPLICANT: GENENTECH, INC.  
TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/03718  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Wendy M. Lee  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 879PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
PCT-US95-03718-4

Query Match 99.4%; Score 5240; DB 5; Length 993;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 987; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPVLLVFSAMIFGTITNODLPVVKVCLINHNKNDSSVGKSSSYPMVSESP 60  
DB 1 MPALARDGGQLPVLVFSAMIFGTITNODLPVVKVCLINHNKNDSSVGKSSSYPMVSESP 60

QY 61 EDLGCALRPOSSGTVYEAADVDSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120  
DB 61 EDLGCALRPOSSGTVYEAADVDSASITLQVLVDAPGNISCLWVFKHSSLNCQPHFDL 120

QY 121 QNRGVVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSIINTLLYTLRRPYFRKMENQD 180  
DB 121 QNRGVVSMVILKMTETQAGEYLLFIQSEATNYTLFTVSIINTLLYTLRRPYFRKMENQD 180

QY 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240  
DB 181 ALVCISESVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE 240

QY 241 CTRLFTIDLNOTPQTTLPQLFKVGEPLWIRCKAVHVNHGFGLTWELNKALEEGNYFEM 300  
DB 241 CTRLFTIDLNOTPQTTLPQLFKVGEPLWIRCKAVHVNHGFGLTWELNKALEEGNYFEM 300

QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTTCSSSKHPQSOSALVTIVGKGFINATNSSEDEYE 360  
DB 301 STYSTNRTMIRILFAFVSSVARNDTGYTTCSSSKHPQSOSALVTIVGKGFINATNSSEDEYE 360

QY 361 IDOYEEFCFSVRKAYPOIRCTWTFSRKSPCEOKGLDNGYSISKFCNKHKQHPGEYIFHA 420  
DB 361 IDOYEEFCFSVRKAYPOIRCTWTFSRKSPCEOKGLDNGYSISKFCNKHKQHPGEYIFHA 420

QY 421 ENDDAQFTKMTLNIRKPOVLAEASASQSCFSDGYPLPSWTWKCKSDKSPNCTEEITE 480  
DB 421 ENDDAQFTKMTLNIRKPOVLAEASASQSCFSDGYPLPSWTWKCKSDKSPNCTEEITE 480  
QY 481 GYVNRKANRKFQGVQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQD 540  
DB 481 GYVNRKANRKFQGVQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQD 540  
QY 541 NISFYATIGVCLLFIVLTLCHYKKQFRYESQLOMVQVGTSSDNEYFYVDFREYED 600  
DB 541 NISFYATIGVCLLFIVLTLCHYKKQFRYESQLOMVQVGTSSDNEYFYVDFREYED 600  
QY 601 LAWEPPRENLEFGKVLGSAFGKVMNATAYGSKTGVSTQVAVKMLKEKADSSEREAALMS 660  
DB 601 LAWEPPRENLEFGKVLGSAFGKVMNATAYGSKTGVSTQVAVKMLKEKADSSEREAALMS 660  
QY 661 ELKMTQLGSHENIVNLLGACTLSGPIYLIFCYCCYDGLLNLYLRSKREKPHRTWTIFKE 720  
DB 661 ELKMTQLGSHENIVNLLGACTLSGPIYLIFCYCCYDGLLNLYLRSKREKPHRTWTIFKE 720  
QY 721 HNFESFYPTQSHPNSSMPGSRREVQIHPDSDQISGLHNSFHSDEIYEYENOKRLEEEEDL 780  
DB 721 HNFESFYPTQSHPNSSMPGSRREVQIHPDSDQISGLHNSFHSDEIYEYENOKRLEEEEDL 780  
QY 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVVVKICDFGLARDIMSDS 840  
DB 781 NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVVVKICDFGLARDIMSDS 840  
QY 841 NYVVRGNARLPVKWMAPELSFEGIYTIKSDVMSYGILLWEIFSLGVNPPYGPVPDANFYK 900  
DB 841 NYVVRGNARLPVKWMAPELSFEGIYTIKSDVMSYGILLWEIFSLGVNPPYGPVPDANFYK 900  
QY 901 LIONGFKMOPPYATIEIYIIMOSWAFDSRKRPSPNLTSLFGLCOLADAEAEAMYNQVOD 960  
DB 901 LIONGFKMOPPYATIEIYIIMOSWAFDSRKRPSPNLTSLFGLCOLADAEAEAMYNQVOD 960  
QY 961 RVSECPHTYQNRPPFSREMDLGLLSFOAQVEDS 993  
DB 961 PVSECPHTYQNRPPFSREMDLGLLSFOAQVEDS 993

RESULT 11

US-08-183-211-2  
Sequence 2, Application US/08183211  
Patent No. 5618709  
GENERAL INFORMATION:  
APPLICANT: Alan M. Gewirtz, Donald Small, Curt I. Civin.  
TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES  
TITLE OF INVENTION: SPECIFIC FOR STK-1 AND METHOD FOR  
TITLE OF INVENTION: INHIBITING EXPRESSION OF THE STK-1 PROTEIN  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEIDEL GONDA LAVORGNA & MONACO  
STREET: Suite 1800, Penn Center Plaza  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: U.S.A.  
ZIP: 19102  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/183,211  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Monaco, Daniel A.  
REGISTRATION NUMBER: 30,480

REFERENCE/DOCKET NUMBER: 3957-15  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (215) 568-8383  
TELEFAX: (215) 568-5549  
TELEX: No. 5618709e  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-183-211-2

Query Match 99.3%; Score 5238; DB 1; Length 993;

Best Local Similarity 99.7%; Pred. NO. 0;

Matches 991; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 MPALADAGTVPVLLVFSAMIFGTITNODLPVTKCVLINHKNDSSVGKSSSYPMVSESP 60  
DB 1 MPALADAGTVPVLLVFSAMIFGTITNODLPVTKCVLINHKNDSSVGKSSSYPMVSESP 60  
QY 61 EDLGCALRPOSSGTVEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120  
DB 61 EDLGCALRPOSSGTVEAAAEVDVSASITLQVLVDAPGNISCLWVFKHSLNCQPHFDL 120  
QY 121 QNRGVVSMVILKMTETQAGEYLLFIQSEATNYTILFTVSTRNTLLYTLRPYFRKMNQD 180  
DB 121 QNRGVVSMVILKMTETQAGEYLLFIQSEATNYTILFTVSTRNTLLYTLRPYFRKMNQD 180  
QY 181 ALVCISESPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
DB 181 ALVCISESPEPIVWVLCDSQGESCKEESPAVVKKEKVLHELFGTDIRCCARNELGRE 240  
QY 241 CTRLFTIDLNOTPQTLPLQFLKVGEPWIRCAVHVNHGFGLTWELNKALEBEGNFEM 300  
DB 241 CTRLFTIDLNOTPQTLPLQFLKVGEPWIRCAVHVNHGFGLTWELNKALEBEGNFEM 300  
QY 301 STYSTNRTMIRILFAVSSVARNDTGYTCTSSSKHPQSALVTIVGKGFINATSSSEDE 360  
DB 301 STYSTNRTMIRILFAVSSVARNDTGYTCTSSSKHPQSALVTIVGKGFINATSSSEDE 360  
QY 361 IDQYEEFCFSVRFKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNHKHQPOEYIFH- 419  
DB 361 IDQYEEFCFSVRFKAYPOIRCTWTFSRKSPCEQKGLDNGYSISKFCNHKHQPOEYIFH- 420  
QY 420 AENDDAQFTKMTLNIRKPOVLAEASASQSCFSDGYPLPSWTWKCKSDKSPNCTEEIT 479  
DB 420 AENDDAQFTKMTLNIRKPOVLAEASASQSCFSDGYPLPSWTWKCKSDKSPNCTEEIT 479  
QY 480 EGVNRRKANRKFQGVQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQ 539  
DB 480 EGVNRRKANRKFQGVQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQ 539  
QY 540 DNISFYATIGVCLLFIVLTLCHYKKQFRYESQLOMVQVGTSSDNEYFYVDFREYED 599  
DB 540 DNISFYATIGVCLLFIVLTLCHYKKQFRYESQLOMVQVGTSSDNEYFYVDFREYED 599  
QY 600 DLKWEFPRENLEFGKVLGSAFGKVMNATAYGSKTGVSTQVAVKMLKEKADSSEREAALM 659  
DB 600 DLKWEFPRENLEFGKVLGSAFGKVMNATAYGSKTGVSTQVAVKMLKEKADSSEREAALM 659  
QY 660 SELKMTQLGSHENIVNLLGACTLSGPIYLIFCYCCYDGLLNLYLRSKREKPHRTWTIFK 719  
DB 660 SELKMTQLGSHENIVNLLGACTLSGPIYLIFCYCCYDGLLNLYLRSKREKPHRTWTIFK 719  
QY 720 EHNFSFYPTQSHPNSSMPGSRREVQIHPDSDQISGLHNSFHSDEIYEYENOKRLEEEED 779  
DB 720 EHNFSFYPTQSHPNSSMPGSRREVQIHPDSDQISGLHNSFHSDEIYEYENOKRLEEEED 779  
QY 780 LNVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVVVKICDFGLARDIMSD 839  
DB 780 LNVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVVVKICDFGLARDIMSD 839  
QY 840 SNYVVRGNARLPVKWMAPELSFEGIYTIKSDVMSYGILLWEIFSLGVNPPYGPVPDANFY 899

Db 840 SNVVRGNARLPVKWAPESLFGVITIKSDVWSYGILLWEIFSLGVNPPGIPVDANFY 899  
QY 900 KLQNGFKMDOPPYATEEYIIIMQSWAFDSRKRPSFPNLTSLGCOLADAEAMQONVD 959  
Db 900 KLQNGFKMDOPPYATEEYIIIMQSWAFDSRKRPSFPNLTSLGCOLADAEAMQONVD 959  
QY 960 GRVSECPHTYQNRPPSREMDLGLLSPQAQVEDS 993  
Db 960 GRVSECPHTYQNRPPSREMDLGLLSPQAQVEDS 993

RESULT 12  
PCT-US95-00176A-2  
; Sequence 2, Application PC/TUS9500176A  
; GENERAL INFORMATION:  
; APPLICANT: Alan M. Gewirtz, Donald Small, Curt I. Civin.  
; TITLE OF INVENTION: ANTISENSE OLIGONUCLEOTIDES  
; TITLE OF INVENTION: SPECIFIC FOR STK-1 AND METHOD FOR INHIBITING EXPRESSION OF THE  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: SEIDEL GONDA LAVORGNA & MONACO  
; STREET: Suite 1800, Penn Center Plaza  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/00176A  
; FILING DATE: 6 January 1995  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/183,211  
; FILING DATE: 14 January 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 3957-14 PC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5349  
; TELEX: None  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 993 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single stranded  
; TOPOLOGY: linear  
PCT-US95-00176A-2

Query Match 99.3%; Score 5238; DB 5; Length 993;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 991; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 MPALARDAGTVPLLVVFSAMIFCTITNQDLPVTKVCLINHKNDSSVGKSSSYPMWSESP 60  
Db 1 MPALARDAGTVPLLVVFSAMIFCTITNQDLPVTKVCLINHKNDSSVGKSSSYPMWSESP 60  
QY 61 EDLGCALRPOSSGTVYEAAREVDVSASITLQVLVDAPGNISCLWFKHSSLNCQPHFDL 120  
Db 61 EDLGCALRPOSSGTVYEAAREVDVSASITLQVLVDAPGNISCLWFKHSSLNCQPHFDL 120  
QY 121 QNRGVVSWILKMTQAGEYLLFIQSEATNYTLFTVSRNTLLYTLRPPYFRKMNQD 180  
Db 121 QNRGVVSWILKMTQAGEYLLFIQSEATNYTLFTVSRNTLLYTLRPPYFRKMNQD 180  
QY 181 ALVCISESPPIVWVLCDSQGESCKEESPAVYKKEKVLHELFGTDIRCCARNELGRE 240

Db 181 ALVCISESPPIVWVLCDSQGESCKEESPAVYKKEKVLHELFGTDIRCCARNELGRE 240  
QY 241 CTRLFTIDLNTQTPOTTLPOLFLKVGEPDLWIRKAVHNHNGGLTWELNKALEGNFYEM 300  
Db 241 CTRLFTIDLNTQTPOTTLPOLFLKVGEPDLWIRKAVHNHNGGLTWELNKALEGNFYEM 300  
QY 301 STYSTNRTMIRILFAFVSSVARNDTGYTTCSSSKHPQSALVTIVGKGFINATNSSEDEY 360  
Db 301 STYSTNRTMIRILFAFVSSVARNDTGYTTCSSSKHPQSALVTIVGKGFINATNSSEDEY 360  
QY 361 IDOYEFCFSVRKAYPOIRCTWTFSTRKSPCEQKGLDNGYSISKFCNKHKHPGEYIFH- 419  
Db 361 IDOYEFCFSVRKAYPOIRCTWTFSTRKSPCEQKGLDNGYSISKFCNKHKHPGEYIFH- 420  
QY 420 AENDDAQFTKMTLNIRKPOVLAEASASQASCFSDGYPPLPSWTWKKCDSPNCTEIT 479  
Db 421 AENDDAQFTKMTLNIRKPOVLAEASASQASCFSDGYPPL-SWTWKKCDSPNCTEIT 479  
QY 480 EGVNRRKANRKFVQWVSSSTLNMSAIGFLVKCAVNSLGTSCETILLNSPGFPPIQ 539  
Db 480 EGVNRRKANRKFVQWVSSSTLNMSAIGFLVKCAVNSLGTSCETILLNSPGFPPIQ 539  
QY 540 DNISFYATIGVCLLFIVVLTLLICHKYKQPRYESQOLQWQVGTSSDNEYFYVDFREY 599  
Db 540 DNISFYATIGVCLLFIVVLTLLICHKYKQPRYESQOLQWQVGTSSDNEYFYVDFREY 599  
QY 600 DLKWEFPRENLEFGKVLGSGAFGVMNATAYGISKTVSIOVAVKMLKEKADSSERELM 659  
Db 600 DLKWEFPRENLEFGKVLGSGAFGVMNATAYGISKTVSIOVAVKMLKEKADSSERELM 659  
QY 660 SELKMTQLGSHENIVNLGACTLSGPTLYLIFYCCYGDLLNLYRSRKRKHRTWTEIFK 719  
Db 660 SELKMTQLGSHENIVNLGACTLSGPTLYLIFYCCYGDLLNLYRSRKRKHRTWTEIFK 719  
QY 720 EHNFSEYPTFQSHENSSMPGSRVQIHPDSQOISLGHNSFHSDEIEYENOKRLEEED 779  
Db 720 EHNFSEYPTFQSHENSSMPGSRVQIHPDSQOISLGHNSFHSDEIEYENOKRLEEED 779  
QY 780 LNVLTFFEDLLCFAYQVAKGMEFFKSCVHRDLAARNVLTGHKVVKICDFGLARDINSD 839  
Db 780 LNVLTFFEDLLCFAYQVAKGMEFFKSCVHRDLAARNVLTGHKVVKICDFGLARDINSD 839  
QY 840 SNVVRGNARLPVKWAPESLFGVITIKSDVWSYGILLWEIFSLGVNPPGIPVDANFY 899  
Db 840 SNVVRGNARLPVKWAPESLFGVITIKSDVWSYGILLWEIFSLGVNPPGIPVDANFY 899  
QY 900 KLQNGFKMDOPPYATEEYIIIMQSWAFDSRKRPSFPNLTSLGCOLADAEAMQONVD 959  
Db 900 KLQNGFKMDOPPYATEEYIIIMQSWAFDSRKRPSFPNLTSLGCOLADAEAMQONVD 959  
QY 960 GRVSECPHTYQNRPPSREMDLGLLSPQAQVEDS 993  
Db 960 GRVSECPHTYQNRPPSREMDLGLLSPQAQVEDS 993

RESULT 13  
US-08-222-299-2  
; Sequence 2, Application US/08222299  
; Patent No. 5635388  
; GENERAL INFORMATION:  
; APPLICANT: Bennett, Brian D.  
; APPLICANT: Broz, Susan D.  
; APPLICANT: Matthews, William  
; APPLICANT: Zeigler, Francis C.  
; TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080



COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/222,299  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 879  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1000 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-222-299-2

Query Match 86.0%; Score 4533.5; DB 1; Length 1000;  
Best Local Similarity 85.6%; Pred. No. 0;  
Matches 851; Conservative 56; Mismatches 84; Indels 3; Gaps 2;

QY 1 MPALA-RDAGTVPLLVFAMIFGTTNODLPVVKVILNHNKNDSSVGKSSSPVWVSES 59  
DB 1 MRAQAQRDRRLLLVLSVMTITVNDLPVVKVILSHENNSSAGKPSIRWVRGS 60  
QY 60 PEDLCALRQSSGTYVEAAAEVDVSASITLQVLDVDPAGNITSLVWFKHSSLCQPHFD 119  
DB 61 PEDLOCTPRQSEGYVEATVEAESGSTITLQVQLATPCDLSCVWFKHSSLCQPHFD 120  
QY 120 LQNRGVVSVILKMTQAGEYLLFTQSEATNTILFTVSRINTLYTLRPFYRKMEQ 179  
DB 121 LQNRGVISMAILNVTQAGEYLLHTQSEAAANYTLFTVNRDQTLVLRPFYRKMEQ 180  
QY 180 DALVCISESVPEPIVEMVLCDSQESCKEESPAVVKKEKVLHFGTDIRCCARNELGR 239  
DB 181 DALLCISEGPEPTVEVWVLCSSHRESCKEEGPAVVRKEEKVLHFGTDIRCCARNALGR 240  
QY 240 ECTRLFTIDNQTPTTLPLQLFLKVGEPWIRKAVVHNGHGLTWELNKALEEGNYFE 299  
DB 241 ESTKLTIDLQAQSTLPLQLFLKVGEPWIRKATHVNHGFLTWELDKALEEGSYFE 300  
QY 300 MSTYSTNRTMIRILFAPVSSVARNDTGYTCCSSKHPSSQSAVLTIVGKGFINATNSSEY 359  
DB 301 MSTYSTNRTMIRILLAFVSSVGRNDGYTCCSSKHPSSQSAVLTILEKGFINATSSQEEY 360  
QY 360 EIDOVEFCFSVRFKAYPOIRCTWTSRKSFPCEQGLDNGYSISKPCNKHOPGEYIFH 419  
DB 361 EIDPYEKFCSVRFKAYPRIRCTWISFQSPCEQGLDNGYSISKPCDKHNPGEYIFY 420  
QY 420 AENDDAQFTKMTNIRRKQVLAESAASQSCFSDGYPLPSTWTKKCDKSPNCTEIT 479  
DB 421 AENDDAQFTKMTNIRKQVLANASASQSCSDGYPLPSTWTKKCDKSPNCTEIP 480  
QY 480 EGVNKRANKRVQGVWSSSTLNMSBAIKFLVKCCAYNSLGTSCETILNLSGPPFFFIQ 539  
DB 481 EGVNKRANKRVQGVWSSSTLNMSBAIKFLVKCCAYNSLGTSCETIFLNSGPPFFFIQ 540  
QY 540 DNISFATIGVCLLFIWLTLLCHYKKKQFRYESQLOMQVQTGSSDNEYFYVDREY 599  
DB 541 DNISFATIGVCLLFIWLTLLCHYKKKQFRYESQLOMQVQTGSSDNEYFYVDREY 600  
QY 600 DLKWEPPRENLEFGKVLGSGAFGRVNMATAYGISKTGVSIOQVAVKMLKEKADSCKEALM 659

DB 601 DLKWEPPRENLEFGKVLGSGAFGRVNMATAYGISKTGVSIOQVAVKMLKEKADSCKEALM 660  
QY 660 SELKMMTQLGSHENIVNLLGACTLSQPIYLIFEYCCYGDLLNLYRSKRKFKHRTWTEIFK 719  
DB 661 SELKMMTHLGHHDNIVNLLGACTLSGPVYLIFEYCCYGDLLNLYRSKRKFKHRTWTEIFK 720  
QY 720 EHNFSYPTFQSHPNSSMPGSRREVQIHPSDDQISGLHGNFSHSEDEIEYENQKRL--EEE 777  
DB 721 EHNFSYPTFQSHPNSSMPGSRREVQIHPSDDQISGLHGNFSHSEDEIEYENQKRLAEEE 780  
QY 778 EDNLVLTFFEDLLCFAYQVAKGMFELEKSCVHRDLAARNVLTGHGVKVICDFGLARDIM 837  
DB 781 EDNLVLTFFEDLLCFAYQVAKGMFELEKSCVHRDLAARNVLTGHGVKVICDFGLARDIL 840  
QY 838 SDSNYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGIILLWEIFSLGVNPNYPGIPVDAN 897  
DB 841 SDSNYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGIILLWEIFSLGVNPNYPGIPVDAN 900  
QY 898 FYKLIONGFKMDQPFYATEEIIYIMOSWAFDSKRKPSFPNLTSLFCCQLADAEAEAMQON 957  
DB 901 FYKLIOGFKMEQPFYATEGIIYFVMSQWAFDSKRKPSFPNLTSLFCCQLADAEAEAMQON 960  
QY 958 VDGRTVSECPHTYQNRPPFSREMDLGLLSPOAOVE 991  
DB 961 MGNVPEHPHSIYQNRRLPSREAGSEPPSPQAVK 994

RESULT 14  
US-08-434-878-2  
Sequence 2, Application US/08434878  
Patent No. 5997865  
GENERAL INFORMATION:  
APPLICANT: Bennett, Brian D.  
APPLICANT: Broz, Susan D.  
APPLICANT: Matthews, William  
APPLICANT: Zeigler, Francis C.  
TITLE OF INVENTION: AGONIST ANTIBODIES AGAINST THE FLK2/FLT3 RECEPTOR AND USES  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/434,878  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Hasak, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 879  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1000 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-434-878-2

Query Match 86.0%; Score 4533.5; DB 2; Length 1000;

[illegible]

RESULT 15

```

PCT-US95-03718-2
: Sequence 2, Application PC/TUS9503718
: GENERAL INFORMATION:
: APPLICANT: GENENTECH, INC.
: TITLE OF INVENTION: AGONIST ANTIBODI
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Genentech, Inc.
: STREET: 460 Point San Bruno Blvd
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 5.25 inch, 360 Kb flo
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: patin (Genentech)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/03718
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Wendy M. Lee
: REGISTRATION/DOCKET NUMBER: 879PCT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415/225-1994
: TELEFAX: 415/952-9881
: TELEX: 910/371-7168
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1000 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
PCT-US95-03718-2

```

Query. Match	86.0%;	Score 4533.5;	DB 5;	Length 1000;
Best Local Similarity	85.6%;	Pred. No. 0;		
Matches	851;	Conservative	56;	Mismatches 84; Indels 3; Gaps 2;
QY	1	MPALA-RDAGTVPLLVFVSAMIFGTFITNODLPVTKCVLINHKNDSSVGKSSSYPMVSES	59	
Db	1	MRALAQRSDRLLLVVLVSMILETVTNODLPVTKCVLISHENNGSSAGKPSSTRMVRGS	60	
QY	60	PEDUGCALRPOSSGTYVEAAAEVDVSASITLQVLVDAPGNISCLWLFKHSLNCPQHFD	119	
Db	61	PEDLQCTPRRQSEGTVEAATVEAEGSITLQVOLATPGDLSCLWLFKHSLSGCQPHFD	120	
QY	120	LQNRGVVSMVLKMTETQAGEYLLFIOSEATNYTILFETVSRNTLLYTLRPPYPRKMEQ	179	
Db	121	LQNRGIYSMAILNVTETQAGEYLHIHIOSEANYITLVTVNVRDTQLVLRPPYPRKMEQ	180	
QY	180	DALVCISESVPEPIVEWVLCDQSQESKEESPAYVKEEVLHLELFGTDIRCCARNELGR	239	
Db	181	DALLCISEGVPEPTVEWVLCDSSHRECKEGPAVVRKEEVLHLELFGTDIRCCARNALGR	240	
QY	240	ECYRLFITDLNQTPTTLPOLFLKVGPELATRCKAVHVNHCFGLTWELENKALEEGNYFE	299	
Db	241	ESTKLFITDLNQAQSTLPOLFLKVGPELIRCKAIHVNHGFGLTWELEOKALEEGSYFE	300	
QY	300	MSYTYSTNRTMIRILFAFVSVVARNDTCYTCSSSKHPQSALVTIVGKGFINATNSSEDY	359	
Db	301	MSYTYSTNRTMIRILLAPVSSVGRNDTCYTCSSSKHPQSALVTILEKGFINATSSQEEY	360	
QY	360	EIDQYEEFCFSVRKAYPQIRCTWTFPSKFPCEQGLDNGYSISKFCNHNKHQGEYIFH	419	
Db	361	EIDPYEKFCSVRKAYPRIRCTWIFSQAFGEQORLEDGYSLSKFCDNHNKHPGEYIFY	420	
QY	420	AENDDAQFTKMTFINIRRRKQVLAESAASQAFSDGYPPLPSMTWTKKCSOKSPNCTBEIT	479	

Db	421	AENDAQTKMFTLNIRKKQVLANASQASCSDDGYPLPSWTWKKCSDRSPNCTEIP	480
Qy	480	EGVWNRKANRVFGQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPFPIQ	539
Db	481	EGVWNRKANRVFGQWSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPFPIQ	540
Qy	540	DNISFYATIGVCLLFIIVVLTLLICHKKQFRIYESOLOQWQVGTSSDNEYFYVDFREY	599
Db	541	DNISFYATIGLCLFIIVVLTLLICHKKQFRIYESOLOQWQVGTSSDNEYFYVDFREY	600
Qy	600	DLKWEFFRENLEFGKVLGSGAFGKVMNATAYGISTGTYSIQVAVKMLKADSSREALM	659
Db	601	DLKWEFFRENLEFGKVLGSGAFGKVMNATAYGISTGTYSIQVAVKMLKADSSREALM	660
Qy	660	SELKMTQLGSHENIVNLGACTLSGPYLIPEYCCYGDLLNLRKSKREKFRHTWTEIFK	719
Db	661	SELKMTQLGSHENIVNLGACTLSGPYLIPEYCCYGDLLNLRKSKREKFRHTWTEIFK	720
Qy	720	EHNFSTYFTFQSHPNSSMPGSRVQIHPDSDOIISGLHNSFHSDEIEYENOKRL--EEE	777
Db	721	EHNFSTYFTFQSHPNSSMPGSRVQIHPDSDOIISGLHNSFHSDEIEYENOKRLAEEEE	780
Qy	778	EDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVVKICDFGLARDIM	837
Db	781	EDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLTGHKVVKICDFGLARDIL	840
Qy	838	SDSNVVRGNARLPVKWMAPELSLFEIGIYTKSDVMSYGILLWEIFSLGVNYPGIPVDAN	897
Db	841	SDSNVVRGNARLPVKWMAPELSLFEIGIYTKSDVMSYGILLWEIFSLGVNYPGIPVDAN	900
Qy	898	FYKLIQNGEKMDOPFYATEEIIYIMOSWAFDSRKRPSFNLTSFLGCOLADAEAEAYON	957
Db	901	FYKLIQNGEKMDOPFYATEEIIYIMOSWAFDSRKRPSFNLTSFLGCOLADAEAEAYON	960
Qy	958	VDGRVSECPHTYQNRPRPSREMDLGLLSPQAOVE	991
Db	961	MCGNVPEHPSIYQNRPRPLSREAGSEPPSPQAOVK	994

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:34:40 ; Search time 24.5123 Seconds  
(without alignments)  
4017.505 Million cell updates/sec

Title: US-09-919-408-4  
Perfect score: 5274  
Sequence: 1 MPALARDAGTVPLLVFSAM.....PFSREMDLGLSPQAQVEDS 993

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5274	100.0	993	10	US-09-919-408-4
2	5274	100.0	993	10	US-09-872-136-4
3	4429.5	84.0	992	10	US-09-919-408-2
4	4429.5	84.0	992	10	US-09-872-136-2
5	1266	24.0	975	9	US-10-192-867-2
6	1251	23.7	976	9	US-10-099-895-32
7	1251	23.7	976	9	US-10-192-867-4
8	1216.5	23.1	972	10	US-09-944-807-10
9	1166.5	22.1	1088	9	US-09-961-403-4
10	1166	22.1	1089	9	US-09-955-363-36
11	1166	22.1	1089	10	US-09-769-987-2
12	1166	22.1	1089	10	US-09-919-497-90
13	1166	22.1	1089	10	US-09-866-510-2
14	1163	22.1	1089	10	US-09-866-510-10
15	1162	22.0	1089	10	US-09-866-510-4
16	1161	22.0	1089	10	US-09-866-510-8
17	1160	22.0	1089	10	US-09-866-510-6
18	1086.5	20.6	1090	10	US-09-866-510-14
19	1086.5	20.6	1106	9	US-09-955-363-2

20	1086.5	20.6	1106	10	US-09-866-510-22
21	1082.5	20.5	1106	10	US-09-866-510-16
22	1081.5	20.5	1106	10	US-09-866-510-20
23	1080.5	20.5	1106	10	US-09-866-510-18
24	990	18.8	1338	9	US-10-059-585-44
25	957	18.1	386	9	US-09-939-833-6
26	957	18.1	386	10	US-09-939-754-6
27	957	18.1	386	10	US-09-939-832-6
28	951.5	18.0	1356	9	US-10-022-939-2
29	951.5	18.0	1356	9	US-10-100-405A-2
30	950.5	18.0	1356	9	US-09-969-037-7
31	934	17.7	1367	10	US-09-766-678-2
32	923	17.5	1367	10	US-09-919-408-6
33	923	17.5	1367	10	US-09-872-136-6
34	907.5	17.2	1363	9	US-09-375-248-19
35	904.5	17.2	1298	10	US-09-982-610-33
36	904.5	17.2	1363	9	US-09-375-248-2
37	902.5	17.1	1368	9	US-10-105-901-34
38	902	17.1	1362	9	US-10-105-901-33
39	880.5	16.7	1363	9	US-10-105-901-32
40	859.5	16.3	367	9	US-09-939-833-9
41	859.5	16.3	367	10	US-09-939-754-9
42	859.5	16.3	367	10	US-09-939-832-9
43	846	16.0	367	9	US-09-939-833-12
44	846	16.0	367	10	US-09-939-754-12
45	846	16.0	367	10	US-09-939-832-12

ALIGNMENTS

RESULT 1  
US-09-919-408-4  
; Sequence 4, Application US/09919408  
; Patent No. US20020072077A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemischka, Ihor R.  
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
; RECEPTORS AND THEIR LIGANDS

NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ImClone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/919,408

FILING DATE: 31-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/977,451

FILING DATE: <Unknown>

APPLICATION NUMBER: US 07/906,397

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: US PCT/US92/05401

FILING DATE: 26-JUN-1992

APPLICATION NUMBER: TW 81102961

FILING DATE: 15-APR-1992

APPLICATION NUMBER: US PCT/US92/02750

FILING DATE: 02-APR-1992

APPLICATION NUMBER: US 07/813,593

FILING DATE: 24-DEC-1991

APPLICATION NUMBER: US 07/793,065

FILING DATE: 15-NOV-1991

APPLICATION NUMBER: US 07/728,913

FILING DATE: 28-JUN-1991

```

:
: APPLICATION NUMBER: US 07/679,666
: FILING DATE: 02-APR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Feit, Irving N.
: REGISTRATION NUMBER: 28,601
: REFERENCE/DOCKET NUMBER: LEM-3-7P
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-645-1405
: TELEFAX: 212-645-2054
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 993 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-919-408-4

Query Match          100.0%; Score 5274; DB 10; Length 993;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

QY 1 MPALARDAGTVPLLVVFSAIFGTITNODLPVVKVNLNHNKNDSSVGKSSSPYPMVSESP 60
Db 1 MPALARDAGTVPLLVVFSAIFGTITNODLPVVKVNLNHNKNDSSVGKSSSPYPMVSESP 60

QY 61 EDLGCALRPQSSGTYEAAAEVDVYSASITLQVLVDAPGNISCLVWFKHSSLNCPHFOL 120
Db 61 EDLGCALRPQSSGTYEAAAEVDVYSASITLQVLVDAPGNISCLVWFKHSSLNCPHFOL 120

QY 121 QNRGVVSVILKMTQAGEYLLFTQSEATNTILFTVSIRNTLLYTLRRPYFRKMENOD 180
Db 121 QNRGVVSVILKMTQAGEYLLFTQSEATNTILFTVSIRNTLLYTLRRPYFRKMENOD 180

QY 181 ALVCISEVPPIVWVLCDGSGCKESPAAVVKKEKVLHELFGTDIRCCARNELGRE 240
Db 181 ALVCISEVPPIVWVLCDGSGCKESPAAVVKKEKVLHELFGTDIRCCARNELGRE 240

QY 241 CTRLFTIDLNPQTTLPLQLFKVGEPLWIRKAVHNVHGFGLTWELNKALEEGNYFEM 300
Db 241 CTRLFTIDLNPQTTLPLQLFKVGEPLWIRKAVHNVHGFGLTWELNKALEEGNYFEM 300

QY 301 SPYSTNRTMIRILFAVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEYD 360
Db 301 SPYSTNRTMIRILFAVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEYD 360

QY 361 IDQEEFCFSVRFKAYPOLRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGYIFHA 420
Db 361 IDQEEFCFSVRFKAYPOLRCTWTFSRKSPCEQKGLDNGYSISKFCNKHQPGYIFHA 420

QY 421 ENDDAQFTKMTLNIRRKPOVLAESAASQSCFSDGYPLPSWTWKKCSKSPNCTEITE 480
Db 421 ENDDAQFTKMTLNIRRKPOVLAESAASQSCFSDGYPLPSWTWKKCSKSPNCTEITE 480

QY 481 GWNRRKANRKFQWVSSSTLNMSAIGFLYKCCAYNSLGTSCETILLNSPGPFPIQD 540
Db 481 GWNRRKANRKFQWVSSSTLNMSAIGFLYKCCAYNSLGTSCETILLNSPGPFPIQD 540

QY 541 NISFYATIGVCLLFIIVVTLILCHYKKQFQRYESQWQVQVTSQSSNEYFYVDFREYD 600
Db 541 NISFYATIGVCLLFIIVVTLILCHYKKQFQRYESQWQVQVTSQSSNEYFYVDFREYD 600

QY 601 LKWEFFRENLEFGKVLGSGAFGKVMNATYVIGSKTGVSIQVAVKMLKEKADSSERIALMS 660
Db 601 LKWEFFRENLEFGKVLGSGAFGKVMNATYVIGSKTGVSIQVAVKMLKEKADSSERIALMS 660

QY 661 ELKMMTQLGSHENIVNLLGACTLSGPIYLIFCYCCYGDLLNLYRSREKFRHTWTEIFKE 720
Db 661 ELKMMTQLGSHENIVNLLGACTLSGPIYLIFCYCCYGDLLNLYRSREKFRHTWTEIFKE 720

QY 721 HNFSEFYPTFQSHNSNMPGSRVQIHPDSDQISGLHGNFSHSEDEIEYENOKRLEEEEDL 780
Db 721 HNFSEFYPTFQSHNSNMPGSRVQIHPDSDQISGLHGNFSHSEDEIEYENOKRLEEEEDL 780

```

```

QY 781 NVLTFEDLLCFAYOVAKGMEFLEFKSCVHRDLAARNVLVTHGKVKVVICDFGLARDIMSDS 840
Db 781 NVLTFEDLLCFAYOVAKGMEFLEFKSCVHRDLAARNVLVTHGKVKVVICDFGLARDIMSDS 840

QY 841 NYVVRGNARLPVKWMAPELSEFEGYITIKSDVMSYGILLWEIPSLGVNPNYPGIPVDANFYK 900
Db 841 NYVVRGNARLPVKWMAPELSEFEGYITIKSDVMSYGILLWEIPSLGVNPNYPGIPVDANFYK 900

QY 901 LQIONFKMDQPFYATEEYIIMQSCWAFDSRRKPSFPNLTSLGCOLADAEAEAMYNQVVG 960
Db 901 LQIONFKMDQPFYATEEYIIMQSCWAFDSRRKPSFPNLTSLGCOLADAEAEAMYNQVVG 960

QY 961 RVSECPHTYQNRPRPSREMDLGLLSPQAQVEDS 993
Db 961 RVSECPHTYQNRPRPSREMDLGLLSPQAQVEDS 993

RESULT 2
US-09-872-136-4
; Sequence 4, Application US/09872136
; Patent No. US20020119545A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/872,136
; FILING DATE: 01-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/208,786
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/09/021,324
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/07/977,451
; FILING DATE: 1992-11-19
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: US PCT/US92/05401
; FILING DATE: 26-JUN-1992
; APPLICATION NUMBER: TW 81102961
; FILING DATE: 15-APR-1992
; APPLICATION NUMBER: US PCT/US92/02750
; FILING DATE: 02-APR-1992
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-7P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054

```

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 993 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-872-136-4

Query Match 100.0%; Score 5274; DB 10; Length 993;  
Best Local Similarity 100.0%; Pred No. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPALARDAGTVPLLVVFSAIFGTTITNODLPVVKVLIHKNKNDSSVKGKSSSPYPMVSESP 60  
Db 1 MPALARDAGTVPLLVVFSAIFGTTITNODLPVVKVLIHKNKNDSSVKGKSSSPYPMVSESP 60  
QY 61 EDLGCALRQSSGTVYAAAEVDVSAITLQVLVDAPGNISCLWVPKHSLSNCQPHFDL 120  
Db 61 EDLGCALRQSSGTVYAAAEVDVSAITLQVLVDAPGNISCLWVPKHSLSNCQPHFDL 120  
QY 121 QNRGVVSWILKMTQAGYLLFIOSEATNYTILFTVSRNTLLYTLRPPYRKMEQD 180  
Db 121 QNRGVVSWILKMTQAGYLLFIOSEATNYTILFTVSRNTLLYTLRPPYRKMEQD 180  
QY 181 ALVCISESPEPIVENVWLCDSQGESCKEESPAVVKKKEKVLHELFGTDIRCCARNELGRE 240  
Db 181 ALVCISESPEPIVENVWLCDSQGESCKEESPAVVKKKEKVLHELFGTDIRCCARNELGRE 240  
QY 241 CTRLFTIDLNTQPTTLPOLFLKVGPELWIRKAVHVNHGFGTLWELNKALEEGNYFEM 300  
Db 241 CTRLFTIDLNTQPTTLPOLFLKVGPELWIRKAVHVNHGFGTLWELNKALEEGNYFEM 300  
QY 301 STYSTNRTMIRILFAVSSVARNDTGYTSCSSKHPQSALVTIVGKGFNATNSSEDE 360  
Db 301 STYSTNRTMIRILFAVSSVARNDTGYTSCSSKHPQSALVTIVGKGFNATNSSEDE 360  
QY 361 IDQEEFCFSVRKAYPOLRCTWTFRRKGFPCQKGLDNGYSISKFCNKHQPGEXIFHA 420  
Db 361 IDQEEFCFSVRKAYPOLRCTWTFRRKGFPCQKGLDNGYSISKFCNKHQPGEXIFHA 420  
QY 421 ENDDAQTFTMFTLNIRRKQVLAESAQSCSDGYPLPSWTWKCSKSPNCTEITE 480  
Db 421 ENDDAQTFTMFTLNIRRKQVLAESAQSCSDGYPLPSWTWKCSKSPNCTEITE 480  
QY 481 GVMNRKANRKFQGVSSSTLNNSAIGKFLVKCCAYNSLGTSCETILLNSPGPFPIQD 540  
Db 481 GVMNRKANRKFQGVSSSTLNNSAIGKFLVKCCAYNSLGTSCETILLNSPGPFPIQD 540  
QY 541 NISFYATIGVCLLFIVLTLILCHIKKQFRYESQLOMVQVGTSSDNEXYFVDFREYED 600  
Db 541 NISFYATIGVCLLFIVLTLILCHIKKQFRYESQLOMVQVGTSSDNEXYFVDFREYED 600  
QY 601 LKWEPPRENLEFGKVLGSAFGKVMNATAYGKTSQVSIQVAVKMLKEKADSSEREA 660  
Db 601 LKWEPPRENLEFGKVLGSAFGKVMNATAYGKTSQVSIQVAVKMLKEKADSSEREA 660  
QY 661 ELKMTQLGSHENIVNLLGACTLSGPIYLIFEYCCYDGLLNLVLRKREKPHRTWTFI 720  
Db 661 ELKMTQLGSHENIVNLLGACTLSGPIYLIFEYCCYDGLLNLVLRKREKPHRTWTFI 720  
QY 721 HNFSEYPTQSHPNSSMPGSRREVQIHPDSDQISGLHNSFHSDEIEYENOKRLEEEEDL 780  
Db 721 HNFSEYPTQSHPNSSMPGSRREVQIHPDSDQISGLHNSFHSDEIEYENOKRLEEEEDL 780  
QY 781 NVLTEDLLCFAYQVAKMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840  
Db 781 NVLTEDLLCFAYQVAKMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDS 840  
QY 841 NYVVRGNARLPKWWAPESLFGIYTIKSDVWSYIGILLWEIFSLGVNPNPPIVDANFYK 900  
Db 841 NYVVRGNARLPKWWAPESLFGIYTIKSDVWSYIGILLWEIFSLGVNPNPPIVDANFYK 900

QY 901 LIQNGFKMDQFPYATEEYIIIMQSCWAFDSRKRPSFPNLTSLFLGCQLADAEAMYNQYVDG 960  
Db 901 LIQNGFKMDQFPYATEEYIIIMQSCWAFDSRKRPSFPNLTSLFLGCQLADAEAMYNQYVDG 960  
QY 961 RVSECPHTYQNRPRPSREMDLGLLSPOAQVEDS 993  
Db 961 RVSECPHTYQNRPRPSREMDLGLLSPOAQVEDS 993

RESULT 3  
US-09-919-408-2  
Sequence 2, Application US/09919408  
Patent No. US20020072077A1  
GENERAL INFORMATION:  
APPLICANT: Lemischka, Ibor R.  
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL  
RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ImClone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/919,408  
FILING DATE: 31-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/977,451  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US 07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US 07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US 07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US 07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US 07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 992 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-919-408-2

Query Match 84.0%; Score 4429.5; DB 10; Length 992;  
Best Local Similarity 84.1%; Pred. No. 1.7e-288;  
Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

1 MPALA-RDAGTVPVLLVFSAMIFGTITNODLPVTKVILNHNKNDSSVSKSSYPMWSES 59  
1 MRALAQRSDRRLLLVLSVMIETVTNODLPVTKVILSHENNSSAGKPSYRMVRGS 60  
60 PEDLCALRQSSGTVYEAFAVEVDVSAITLQVLVDAPGNISCLWVFKHSSLNQCOPHD 119  
61 PEDLOCTPRQSGTVYEAFAVEVDVSAITLQVLVDAPGNISCLWVFKHSSLNQCOPHD 120  
120 LQNRGVVSWILKMTETQAGEYLLFIQSEATNYTLFTVSVIRNTLLYLRPRYFRKMNQ 179  
121 LQNRGVVSWILKMTETQAGEYLLFIQSEATNYTLFTVSVIRNTLLYLRPRYFRKMNQ 180  
180 DALVCISSEVPEPIVWVLCDSQSGESKPEPAVVKKEKVLHELFGTDIRCARNELGR 239  
181 DALCISSEVPEPIVWVLCDSQSGESKPEPAVVKKEKVLHELFGTDIRCARNELGR 240  
240 ECTRLFTIDLNTQPTLPLQFLKVGEPILWIRKAVHNVHNGFGLTWELENKALEGNYFE 299  
241 ECTRLFTIDLNTQPTLPLQFLKVGEPILWIRKAVHNVHNGFGLTWELENKALEGNYFE 300  
300 MSTYSTNTRMIRILFAFVSSVARNDTGYTSCSSKHPQSALVTIVGKFINATNSSEYD 359  
301 MSTYSTNTRMIRILFAFVSSVARNDTGYTSCSSKHPQSALVTIVGKFINATNSSEYD 360  
360 EIDQVEEFCFVRKAYPOIRCTWTFSRKSECEOKGLDNGYSISKFCNKHKPGGEYIFH 419  
361 EIDQVEEFCFVRKAYPOIRCTWTFSRKSECEOKGLDNGYSISKFCNKHKPGGEYIFH 420  
420 AENDDAQETKMTLNRKPKQVLAASASQSCFSDGYPLPSWTWKKCKSDKSPNCTEIT 479  
421 AENDDAQETKMTLNRKPKQVLAASASQSCFSDGYPLPSWTWKKCKSDKSPNCTEIT 480  
480 EGVWNRKANRKYFGOWVSSSTLNMSEAIKGLVKCAYNLSLGTSCETILLNSPGPPFTQ 539  
481 EGVWNRKANRKYFGOWVSSSTLNMSEAIKGLVKCAYNLSLGTSCETILLNSPGPPFTQ 540  
540 DNISFYATIGVCLFTVLTLLICHYKQFYESQLOVMQVGTSSDNEFYVDFREY 599  
541 DNISFYATIGVCLFTVLTLLICHYKQFYESQLOVMQVGTSSDNEFYVDFREY 600  
600 DLKWEFPRENLEFGKVLGSGAFKGMVNATAYGKSTGVSQVAVKMLKADSKESREALM 659  
601 DLKWEFPRENLEFGKVLGSGAFKGMVNATAYGKSTGVSQVAVKMLKADSKESREALM 660  
660 SELKMTQLGSHENIVNLLGACTLSGPVILFEYCCYDGLLNLVLRKREKHFHTWTEIFK 719  
661 SELKMTQLGSHENIVNLLGACTLSGPVILFEYCCYDGLLNLVLRKREKHFHTWTEIFK 720  
720 EHNFSYPTFQSHPNSSMPGSRVOIHPDSDQISGLHNGSHSEDETEYENOKRL--EEE 777  
721 EHNFSYPTFQSHPNSSMPGSRVOIHPDSDQISGLHNGSHSEDETEYENOKRL--EEE 780  
778 EDNLVLTFFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLTGHGVKVKTCDFGLARDIM 837  
781 EDNLVLTFFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLTGHGVKVKTCDFGLARDIM 840  
838 SDSNYVVRGNARLPVKWMAPELFEYITIKSDVWSYGILLWEIFSLGVNPNYPCIPVDAN 897  
841 SDSNYVVRGNARLPVKWMAPELFEYITIKSDVWSYGILLWEIFSLGVNPNYPCIPVDAN 900  
898 FYKLIONGKMDOPFFATEYIYIMOSCAWFDKRRKPSFNLTSFLGCQLADAEAEAYQN 957  
901 FYKLIONGKMDOPFFATEYIYIMOSCAWFDKRRKPSFNLTSFLGCQLADAEAEAYQN 957  
958 VQGRVSECHPTYQNRPRFSEMDLGLLSPOAQVE 991  
958 ----IRTSIHLPKQAAPQRG-GLRAQSPQROVK 986

RESULT 4

US-09-872-136-2

; Sequence 2, Application US/09872136

; Patent No. US20020119545A1

GENERAL INFORMATION:  
APPLICANT: Lemischka, Ihor R.  
TITLE OF INVENTION: TOPIPOTENT HEMATOPOIETIC STEM CELL  
RECEPTORS AND THEIR LIGANDS  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ImClone Systems Incorporated  
STREET: 180 Varick Street  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10014  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/872,136  
FILING DATE: 01-Jun-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/09/208,786  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/09/021,324  
FILING DATE: <Unknown>  
APPLICATION NUMBER: US/07/977,451  
FILING DATE: 1992-11-19  
APPLICATION NUMBER: US/07/906,397  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: US PCT/US92/05401  
FILING DATE: 26-JUN-1992  
APPLICATION NUMBER: TW 81102961  
FILING DATE: 15-APR-1992  
APPLICATION NUMBER: US PCT/US92/02750  
FILING DATE: 02-APR-1992  
APPLICATION NUMBER: US/07/813,593  
FILING DATE: 24-DEC-1991  
APPLICATION NUMBER: US/07/793,065  
FILING DATE: 15-NOV-1991  
APPLICATION NUMBER: US/07/728,913  
FILING DATE: 28-JUN-1991  
APPLICATION NUMBER: US/07/679,666  
FILING DATE: 02-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Felt, Irving N.  
REGISTRATION NUMBER: 28,601  
REFERENCE/DOCKET NUMBER: LEM-3-7P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-645-1405  
TELEFAX: 212-645-2054  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 992 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-872-136-2

Query Match 84.0%; Score 4429.5; DB 10; Length 992;  
Best Local Similarity 84.1%; Pred. No. 1.7e-288;  
Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;

QY 1 MPALA-RDAGTVPVLLVFSAMIFGTITNODLPVTKVILNHNKNDSSVSKSSYPMWSES 59  
Db 1 MRALAQRSDRRLLLVLSVMIETVTNODLPVTKVILSHENNSSAGKPSYRMVRGS 60  
QY 60. PEDLCALRQSSGTVYEAFAVEVDVSAITLQVLVDAPGNISCLWVFKHSSLNQCOPHD 119  
Db 61. PEDLOCTPRQSGTVYEAFAVEVDVSAITLQVLVDAPGNISCLWVFKHSSLNQCOPHD 120  
QY 120 LQNRGVVSWILKMTETQAGEYLLFIQSEATNYTLFTVSVIRNTLLYLRPRYFRKMNQ 179



```
121 LQNRGIVSMAILNVTQAGEYLLHIQSERANTVLTFTVNRDTQLVLRPFRKMEQ 180
180 DALVCISEVPEDIVWVLCDSOGESCKESPAPVKEEVHLHFLGTDTRCCARNLGR 239
181 DALLCISEGPEPTVWVLCSSRESCKEPAPVKEEVHLHFLGTDTRCCARNLGR 240
240 ECTRLFTIDLNQPTTLPLFLKVGEPWIRCKAVHNVHFGTLWELENKALEEGNYFE 299
241 ECTKLTIDLNQAPQSTLPFLKVGEPWIRCKAVHNVHFGTLWELENKALEEGNYFE 300
300 MSTYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEY 359
301 MSTYSTNRTMIRILFAFVSSVGRNDTGYTCSSSKHPSQSALVTILEKGFINATSSQEEY 360
360 EIDOEYEFCSVRPKAYPQIRCTWTFSRKFPCQEQKGLDNGYSISKPCNKHKHPGEYIFH 419
361 EIDPYEKFCFSVRPKAYPQIRCTWTFSRKFPCQEQKGLDNGYSISKPCNKHPGEYIFY 420
420 AENDDAQFTKMTFLNIRRRKPOVLAESAASQACFSQGYPLPSPWTWKKCSKSPNCTEIT 479
421 AENDDAQFTKMTFLNIRRRKPOVLAESAASQACSSQGYPLPSPWTWKKCSKSPNCTEIP 480
480 EGVNRRKANRKFVQGVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQ 539
481 EGVNRRKANRKFVQGVSSSTLNMSAIGFLVKCCAYNSLGTSCETIFLNSPGPPFIQ 540
540 DNISFYATTGCVLLFIVLTLCHYKKGFYRESQLOMVOVTGSSDNEYFYVDFREY 599
541 DNISFYATTGCVLLFIVLTLCHYKKGFYRESQLOMVOVTGSSDNEYFYVDFREY 600
600 DLKWEPPRENLEFGKVLGSGAFKGVNATAYGISKTCVSIQVAVKMLKADSSERALM 659
601 DLKWEPPRENLEFGKVLGSGAFKGVNATAYGISKTCVSIQVAVKMLKADSSERALM 660
660 SELKMMTQGLSHENIYNLLGACLTSGPIYLFYCCYGGDLNLVLRKREKFRHTWTIEFK 719
661 SELKMMTQGLSHENIYNLLGACLTSGPIYLFYCCYGGDLNLVLRKREKFRHTWTIEFK 720
720 EHNFSYPTFQSHNPSMPSRQVTHPSDQISGLHNSFHSDEIEYENQKRL--EEE 777
721 EHNFSYPTFQSHNPSMPSRQVTHPSDQISGLHNSFHSDEIEYENQKRLABEEE 780
778 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIM 837
781 EDNLVLTFFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIL 840
838 SDSNYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 897
841 SDSNYVVRGNARLPVKWMAPESLFEGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 900
898 FYKLIONGFKMDOPFYATBEIYIIMOSWAFDSRKRPSFNLISFGLCOLADAEAMQON 957
901 FYKLIONGFKMDOPFYATBEIYIIMOSWAFDSRKRPSFNLISFGLCOLADAEAMQON 957
958 VDRGVSECPHTYQNRPRFSPREMDGLLSPOAQVE 991
958 ----ITSHLPKQAPQRG--GLRAQSFQK 986
```

```
RESULT 5
US-10-192-867-2
; Sequence 2, Application US/10192867
; Publication No. US20030084466A1
; GENERAL INFORMATION:
; APPLICANT: BLUME-JENSEN, Peter
; APPLICANT: HUNTER, Tony
; TITLE OF INVENTION: SYSTEM AND METHOD FOR CONTROLLING MALE FERTILITY
; FILE REFERENCE: SALKINS.002C1
; CURRENT APPLICATION NUMBER: US/10/192,867
; CURRENT FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: US 60/175,625
; 2000-01-11
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00573
; 2001-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 975
; TYPE: PRT
; ORGANISM: MUS MUSCULUS
US-10-192-867-2
```

```
Query Match 24.0%; Score 1266; DB 9; Length 975;
Best Local Similarity 33.2%; Pred. No. 1.2e-76;
Matches 336; Conservative 173; Mismatches 330; Indels 172; Gaps 41;
```

```
QY 35 CVLINHNNDSSVCKSSYPMVSESPDLGCLALPQSSGTYVEAAAVDVDSASITLQVL 94
DB 12 CVLLVLRGQTATSQPSASGPSP-----SIHPAQ-----ELIVEAGDTL--- 54
QY 95 VDAPGNISCL-----VWFKHSSNLCPHFQDLQNRGVSVILKMTQAGEYLLFIOSE 148
DB 55 -----SLTCDPDPVRWTFK-----TYFN-----EMVENKKNEWIO--EKAE 89
QY 149 ATNYTILFTVSIIRNTL-----LYTLRRPYPRKMNQDALVCISESVPEPIVE 195
DB 90 ATR-TGYTCSNSGLTSSIIYFVRDPAKFLVLGLPLFGK-EDSDALVRCPLTDQ-Q-VSN 146
QY 196 WVLCDOSGESCKEESPAY-----VKKEKVLHLEPGTDIRCCARNELGRECTRL--- 244
DB 147 YSLTECDGKSLPTDLTFVPNPKAGITIKNVRAYHRLC---VRCAAQ-----RDGTWLSHD 199
QY 245 -FTIDLNQOT---PQTTLPOL--FLKYGEPLWIRCKAVHNVHFGELTW-----E 286
DB 200 KFTLVREALKAIIPVSVSPETSHLLKKGDTFTVCTIKDSTSVNSMWLKNPQOQIAQ 259
QY 287 LENKALEEGNYFEMSTYSTRMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVT--- 343
DB 260 VKHNSWRHGD---NVERQETLI-----ISSARVDSDGVPMCYANNTFGSANVTTLK 309
QY 344 IVGKGFTNATN--SEDEYEDQYEFCSVRPKAYPQ--IRCTWTFSRKSFPCEQKGLD--- 398
DB 310 VVEKGFINIIPVKNTTVFTVDGENVDLVVEYEAYPEKHQOQIYMNRT--SANKGKYVK 367
QY 399 -NGYSISKFCNKHK-----QPGYIFHAENDDAQFTKMTFLNIRRRKPOVLA--EASAS 448
DB 368 SDNKSNTIRYNQLRLTRKGTGGTYTFLVNSDASASVTENVYNTKPEILTVDRLNG 427
QY 449 QACFSQGYPLPSPWTWKKCSKSPNCTEITE-EGVNRKANRKFVQGVSSSTLNMSAI 507
DB 428 MLQCVABEGPEPTIDWYFCITGAEQRCITPVPDVOVQVNSVSPGKLVQSSIDSSYFR 487
QY 508 KGLVLCAYNSLGTSCETILLNSPGPPF---IQDN-----ISFYATTGCVLLFTIV 557
DB 488 HNGTVECKASNDVGKS--SAFFN---FAFKEQIQAHFTLPTLLIGFYVAAAGMIIVNV 541
QY 558 LTLILCHYKKQFYRESQLOMV--OVTGSSDNEYFYVDFREYEDLKWEFFRENLEFGKVL 616
DB 542 LT-----YKYLQKPMYEQWKVVEEING---NNYIIDPTQLPDYHKWEEFPRNRLSFGKTL 594
QY 617 GSGAFKGVNATAYGISKTGVSIQVAVKMLKEKADSSERALMSLKMNTQLGSHENIYN 676
DB 595 GAGAFKGVNATAYGLIKSDAANTVAVKMLKPSAHLTEREALMSLKVLSYLGHNMINIVN 654
QY 677 LLGACTLGGPIYILFECYCCGDLNLVLRKREKF-----HRTWTEIFKEHNFIFYPTQ 730
DB 655 LLGACTVGGPTLVITIEYCCYGGDLNLVLRKRDSEIFSKQEBQEAALYKLLHSTEPSCD 714
QY 731 SHPNSSM--PGSREVOIHPSDQISGLHNSFHSDEIEYENQKRLKEEEDLNLVLTED 787
DB 715 S-SNEYMDMKPGVSIV-VPTKTKRRRSARIDSY-----IERDVTPAIMEDDEL-ALDIDD 766
QY 788 LLCFAVOYAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARDIMSDSNVVRGN 847
DB 767 LLESYQVAKMAFLASKNCHIRDLAARNILLTHGRITKICDFGLARDIRDSNVVVRGN 826
```

```
QY 848 ARLPVKNWAPESLFGIYTIKSDVNSYGILLWEIFSLGVPYGPVVDANFYKLIQNGFK 907
DB 827 ARLPVKNWAPESIFSCVYTFSDVNSYGILLWEIFSLGSPYGPMPVDSKFKYMKKEGFR 886
QY 908 MDQFYATEEIIYIMOSCAFWSRFRSPNLTSLFGCOLADAEAMYQNV 958
DB 887 MVSPEHAPAEYDMVKTCHWDADPLKRPTEFKVQVQLIEKISDSTKHIYSNL 937

RESULT 6
US-10-099-895-32
; Sequence 32, Application US/10099895
; Patent No. US20020177166A1
; GENERAL INFORMATION:
; APPLICANT: BERNDT, Michael C
; APPLICANT: STOMSKI, Frank C
; APPLICANT: LOPEZ, Angel F
; APPLICANT: GUTHRIE, Mark A
; TITLE OF INVENTION: A BINDING MOTIF OF A RECEPTOR
; FILE REFERENCE: 3991/0K379
; CURRENT APPLICATION NUMBER: US/10/099,895
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: PCT/AU00/01118
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-099-895-32
```

```
Query Match 23.7%; Score 1251; DB 9; Length 976;
Best Local Similarity 32.7%; Pred. No. 1.2e-75;
Matches 324; Conservative 166; Mismatches 352; Indels 148; Gaps 34;
```

```
QY 47 VGKSSYPMVSESPEDLGCALRPOSSGTVEAAAEVVDVSASITLQVLVDAPGNISCLW 106
DB 20 VQTGSSQPSVSPG-EPSPPSIHPGKSDLI-----VRVGEIRLLCTDPPGFVK--WT 67
QY 107 FKHSLLNCQPHFDLQNRGVVSMVILKMTETQAGEYLLFIQSEATN---YTLFTVSIRNT 163
DB 68 FE-----ILDETENKQNEWIT-EKAEATNTGKYCTCNKHGLSNS 106
QY 164 L-----LYTLRRPYFRKMNQDALVCISESPPEIVWVLCDSQGESCKEE----- 209
DB 107 IYFVRDPKFLVDRSLYLGK-EDNDTLVRCPLTDPE-VTNYSLKGCQGRPLPKDLRFIP 164
QY 210 ---SPAVVKKEKVLHFLGTDIRCCARNELGRECTRLFTIDL---NQTPQTTLPQ--L 260
DB 165 DPKAGIMIKSVKRAYHRLC---LHCSVDQEGKSVLSEKFLKVRPAPKAVPVVSVSKASY 221
QY 261 FLKVGPELWIRCAVHVNHGFLTWELN---KALEEGNYFEMSTYSTNRTMIRILFAFV 317
DB 222 LLREGEFTVTCIKDYSVSVTWKRENSOTKLOEKYNSWHHGFNYEROAT---LTI 277
QY 318 SSVARNDTGYTTCSSSKHPQSALVT---IVGKGFINA-----TNSSEYEDIOYE 365
DB 278 SSARVNDGVMFCYANNTFGSANVTTLLEVVDKGFINFPMINTVFVNDGENDVL----- 333
QY 366 EFCESVFKAYPO-IRCTWTFSRKSF--PCEQKGLDNGYSISKFCNKKH-----OPGE 415
DB 334 ----IVEYEAFPPKEHOQWYIMNFTFDKWEDYPKSESNIRYVSELHLRLKGTGEGT 389
QY 416 YIPHAENDDAQFTKMFNLNIRRKQVOLA--EASASQASCFSDGYPPLPSWTWKKSCKSPN 473
DB 390 YTLVNSVDVNAALAFNVYNTKPEILTYDLVNGMLQCVAAAGPEPIDWYFCPGTEQR 449
QY 474 CTEEBITE-GVNNRKNRKFQWQVSSSTLNMSAEIKGFLVKCCAYNSLGTSCETILLNSP 532
DB 450 CSASVLPVDQTLNNSGPPFGKLVVOSSIDSSAFKHNGTVECKAYNDVGKT--SAYFN-- 505
```

```
QY 533 GPFPFIQDN-----ISFYATIGVCLLFIVVLTLLICHKKYKOPRYESQLOM 578
DB 506 --FAFGNNKEQIHPTLFTPLLLIGFVIVAGMCIILMILT-----YKYLQKPMYEQWKV 559
QY 579 V-QVTGSSDNEFYDFREYEDLKWEPFRENLEFCVGLGSGAFGKVMNATAYGISKTGV 637
DB 560 VBEING---NNYVIDPTQLPYDHDKWEFPNRNLSFKTLGAGAFGVVEATAYGLIKSDA 616
QY 638 STQVAVKMLKEKADSSERREALMSELKMTQOLGSHENIVNLLGACTLSGGPIYLIFECYCG 697
DB 617 AMTVAVKMLKPSAHLTEREALMSELKAVLSYLGNNHMINVLLGACTIGGPTLVITEYCCYG 676
QY 698 DLLNLYRSKREKFRHTWTE-----IFKEHNFSFYTFQSHPNSSM---PGSREVQIHPD 748
DB 677 DLLNFLRRKRDSFICKQEDHAEALYKNLLHKSCESSDSTNEYMDKPGSVV-VPTK 735
QY 749 SDOISGLHGNSFHSEDEIEYENQKRLEEBEDLVNLFEDLLCFAYOVAKGMELEFKSCV 808
DB 736 ADKRRSVRIGSY-----IERDVTTPAIMEDEL-ALDLEDLSFSYQVAKGMAFASKNCI 789
QY 809 HRDLAARNVLVTHGKVVKICDGLARDIMSDSNVVRGNARLPVKWMAPELFEIYTIK 868
DB 790 HRDLAARNILLTHGRITKICDGLARDIKNDSNVYKGNARLPVKWMAPEIFNCVYTFE 849
QY 869 SDVMSYGILLWEIFSLGVPYGPVVDANFYKLIQNGFKMDQPPYATEEIIYIMOSCAF 928
DB 850 SDVMSYGILWELFSLGSSPYGMPVDSKFKYMKIKEGFRMLSPHAPAEYDINKTCWDA 909
QY 929 DSKRPSFNPMTSLFGCOLADAEAMYQNV 958
DB 910 DPLKRPTEFKQIVOLIEKQISESTNHIYSNL 939
```

## RESULT 7

```
US-10-192-867-4
; Sequence 4, Application US/10192867
; Publication No. US2003008466A1
; GENERAL INFORMATION:
; APPLICANT: BLUME-JENSEN, Peter
; APPLICANT: HUNTER, Tony
; TITLE OF INVENTION: SYSTEM AND METHOD FOR CONTROLLING MALE FERTILITY
; FILE REFERENCE: SALKINS-002C1
; CURRENT APPLICATION NUMBER: US/10/192,867
; CURRENT FILING DATE: 2002-07-08
; PRIOR APPLICATION NUMBER: US 60/175,625
; 2000-01-11
; PRIOR APPLICATION NUMBER: PCT/US01/00573
; 2001-01-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 976
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-10-192-867-4
```

Query Match 23.7%; Score 1251; DB 9; Length 976;

Best Local Similarity 32.7%; Pred. No. 1.2e-75;  
Matches 324; Conservative 166; Mismatches 352; Indels 148; Gaps 34;

```
QY 47 VGKSSYPMVSESPEDLGCALRPOSSGTVEAAAEVVDVSASITLQVLVDAPGNISCLW 106
DB 20 VQTGSSQPSVSPG-EPSPPSIHPGKSDLI-----VRVGEIRLLCTDPPGFVK--WT 67
QY 107 FKHSLLNCQPHFDLQNRGVVSMVILKMTETQAGEYLLFIQSEATN---YTLFTVSIRNT 163
DB 68 FE-----ILDETENKQNEWIT-EKAEATNTGKYCTCNKHGLSNS 106
QY 164 L-----LYTLRRPYFRKMNQDALVCISESPPEIVWVLCDSQGESCKEE----- 209
DB 107 IYFVRDPKFLVDRSLYLGK-EDNDTLVRCPLTDPE-VTNYSLKGCQGRPLPKDLRFIP 164
QY 210 ---SPAVVKKEKVLHFLGTDIRCCARNELGRECTRLFTIDL---NQTPQTTLPQ--L 260
```

```
Db 165 DPKAGIMIKSVKRAYHRLC---LHCSVDQEGKSVLSEKFIKVRPAFKAVPVVSVKRSASY 221
Qy 261 FLKVGPEPLATRCRAVHVNHGFGLTWELN---KALEBGNFYFEMSTYNTMTIRILFAFV 317
Db 222 LLREGGEFTVTCIKDVSSSVYTWKRENSQTKLQEKYNSWHHGDFNYERQAT---LTI 277
Qy 318 SSVARNDTGYTTCSSSKHPQSALVT---IVGKGFINA-----TNSSEYEDIOYE 365
Db 278 SSVARVDSGVEMCYANNTEGSAVNTITLEVVDKGFINFIPMINTVTVFVNDGENVDL--- 333
Qy 366 EFCFSVRFKAYPO-ICTWTFSKSP--PCEQKGLONGYSISKFCNKH-----OPGE 415
Db 334----IVEYBAFPKQHQMIYMTTETDKWEDYPKSESNIRYVSELHLTRLKGTGEGT 389
Qy 416 YIFHAENDDAQTKMTLNIARRPOVLA--EASASQASCFSDGYPPLSPWTWKKCSKSPN 473
Db 390 YTLFVNSDVNAIAFNVYNTKPEILTYDRLVNGMLQCVAAAGFPEPTIDWYFCPGTEQR 449
Qy 474 CTEEITE-GVMNRKANRKFVGQWSSSTLNMSEAIKGFVLCCKAYNSLGTSCETILLNSP 532
Db 450 CSASVLPVDQVTLNSSGPFGLVVOSSIDSSAFKHNGTVECKAYNDVGT--SAYFN-- 505
Qy 533 GPFPFTQDN-----ISFYATIGVCLLFTIWLTLILCHIKYKQFYESOLQM 578
Db 506--FAFGNNKEQIHPHTLFTPLLIGFVIVAGMCIIMILT---YKYLQKPMVEQWKV 559
Qy 579 V-QVTGSSDNEFYVDFREYEDLKWEPFRENLEFGKVLGSGAGFKGMNATAYGISTGV 637
Db 560 VEEING---NNVYIIDPTQLPYDHKWEFPNRNLSFGKTLGAGAGFKVVEATAYGLIKSDA 616
Qy 638 STQVAVMKLEKADSSERALMSLKMWTOLGSHENIVNLLGACTLSGPIYLIIFYCCYG 697
Db 617 AMTVAVMKLPKSAHLTERCALMSLKVSLYLGNNHMINVLLGACTTGGPPLVITEYCCYG 676
Qy 698 DLLNLYRSKREKFRHTWTE-----IFKEHNFSEYPTFQSHPNSSM---PGSREVOIHPD 748
Db 677 DLLNLYRRKRDSEFICQEDHAEALYKLLHSEKSSCDSTNEYMDKPGVSYV-VPTK 735
Qy 749 SDQISGLHNSHSEDEIEYENOKRLEEBEDLVLFEDLLCFAYQVAKGMEFLFKSCV 808
Db 736 ADKRSRVISGSY-----IERDVTYPAIMEDEL-ALDLEDLLSFSYQVAKGMFLASKNCI 789
Qy 809 HRDLAARNVLTGHKVVKICDFGLARDIMSDSNVYVGRNARLPVKWMAPELFEIGYTIK 868
Db 790 HRDLAARNLTLLHGRITKICDFGLARDIKNDSNVYVGRNARLPVKWMAPELFEIGYTIK 849
Qy 869 SDWWSYGILLWEIFSLGVNYPGIPVDANFYKLIQNGFKMDQPFYATEEYIIIMQSCWAF 928
Db 850 SDWWSYGIFLWELFSLGSSPYPCMPVDSKRYKMIKEGFRMLSPHAPAEYDIMKTCWDA 909
Qy 929 DSKRPSFNLTSLFCCQLADAEAEANYQV 958
Db 910 DPLKRPTEQIVOLIEKQISESTNHIYSNL 939
```

## RESULT 8

```
US-09-944-807-10
; Sequence 10, Application US/09944807
; Patent No. US2002011949A1
; GENERAL INFORMATION:
; APPLICANT: Boehringer Ingelheim Pharma KG
; TITLE OF INVENTION: Method for identifying substances which positively
; TITLE OF INVENTION: Influence inflammatory conditions of chronic
; TITLE OF INVENTION: Inflammatory airway diseases
; FILE REFERENCE: 082.00n
; CURRENT APPLICATION NUMBER: US/09/944,807
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: UK 0021484.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
```

```
; LENGTH: 972
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-944-807-10

Query Match 23.1%; Score 1216.5; DB 10; Length 972;
Best Local Similarity 32.8%; Pred. No. 2.5e-73;
Matches 325; Conservative 156; Mismatches 316; Indels 195; Gaps 33;

Qy 75 VYEAAYVEVDV--SASITLQVLVDAPGNISCLWFKHSSLNCQPHFDLQNRGVSWVILK 132
Db 22 VIEPSELPVVKQATVTRCV---GNGSVEMDGPSS---PHWTLYSDGSSILSTLN 72
Qy 133 MTE-----TCAGEVL-----LFIOSEATNTYTLFTVSTIRNTLLYTLRRRYEKKM 176
Db 73 NATFQNTGTYRCTEPGDPGLGSSAAIHLVVKDAPRWNVL---AOEVVVF----- 118
Qy 177 ENQDALV-CISEYPEPIVEVWLCDQSGESCKEESPAVVKKEK-----VLHE--- 223
Db 119 EDQDALLPCL---LTDPVLE-----AGVSLVRVGRPLMRHTNYSFSPWHGFTIHRAKF 169
Qy 224 LFGTDIRCCARNELGRECTRLFTIDLNOTPOTTLP-----QLFLKYGEPLWIRCK 273
Db 170 IQSDYQCSALMG-GR---KVMISIRLKVQKVIPGPPALTLVPAELVRIRGEAAQIVCS 225
Qy 274 AVVHNHGFLTWELKENKALEBGNFYFEMSTYSTNRTMIRILFAFYSSVARNDTYTCCSS 333
Db 226 ASSVDVNFDFVLQHNNTKLA---IPOOSDFHNNRYQ-KVLTLNLDQVDFQHAGNYSVAS 281
Qy 334---KHPSSQASLVITVGKGFINATNSSD---YEIDQVEEFCFSVRFKAYPOIR-CTWTF 385
Db 282 NVQCKH-STSMFFRVVESAYLNL--SSEONLQIOWTVGEGNLKLVWEAYPGLQGFNTY 338
Qy 386 SRKSFCEQKGLDNGYSISKFNHQP-----GEYIF 418
Db 339-----LGPPSDHOPEPKLANATTKDTYRHTFTLSLPLKPAEAGRTSF 381
Qy 419 HAENDDAQTKMFTLNIRRRPOVLAESAASQAS---CFSDGYPLPSTWTKKCSKSPNC 474
Db 382 LARPPGWRALTFTLTRYPEPVSIVTFTNGSGTLLCAASGYQPPNVTWLOCSGHTORC 441
Qy 475 TEEITEGVWN---RKANRKFVGOWSSSTENMSEAIKGFVLCCKAYNSLGTSCETILLN 530
Db 442 DEAQVLQVMDPPYEVLSQEPFHKVTQSLTTLVETLEHNTQYECRAHNSVSGSHAFPI 501
Qy 531 SPGPPFIQONISFYATIGVC---LLFIVTLTLLCHIKYKQFYESOLQMVQVTSDDN 587
Db 502 SAGAHTHPDDEFLTPTVVVACMSIMALLLLLLLLLYKKYKQKPKYQVRWKIIE---SYEGN 559
Qy 588 EYFYVDFREYEDLKWEPFRENLEFGKVLGSGAGFKGMNATAYGISTKTVSIOVAVMKL 647
Db 560 SYTFIDPTQLPTQYNEKWEFPNRNNOFGKTLGAGAFKVVYEAATAFGLKEDAVLKVAVMKL 619
Qy 648 EKADSSEREALMSLKMWTOLGSHENIVNLLGACTLSGPIYLIIFYCCYGLLNLYRSKR 707
Db 620 STAHADEKALMSLKMWTOLGSHENIVNLLGACTHGGPVLVITEYCCYGLLNLYRSKR 679
Qy 708 E-----KFHRTWTEIFKEHNFSEYPTFQSHPNSSMPCGSREVOIHP 747
Db 680 EAMLGPSLSPQDPEGGVDYKNIHLEKKYVRRDSGFS-----SQGVDIYV 724
Qy 748 DSQOISGLHNSHSEDEIEYENOKRLEEBEDLVLFEDLLCFAYQVAKGMEFLFKSC 807
Db 725 EMRPVSTSSNDSF-SEQDLKEDGRPLE-----LRDLHFSSQVQAQGMFLASKNC 774
Qy 808 VHRDLAARNVLTGHKVVKICDFGLARDIMSDSNVYVGRNARLPVKWMAPELFEIGYTI 867
Db 775 IHRDVAARNVLLTNGHVAKIGDFGLARDIMSDSNVYVGRNARLPVKWMAPELFEIGYTI 834
Qy 868 KSDWWSYGILLWEIFSLGVNYPGIPVDANFYKLIQNGFKMDQPFYATEEYIIIMQSCWA 927
Db 835 QSDWWSYGILLWEIFSLGVNYPGIPVILYNSKIFYKLVDGQYQMNQPAFAPKNIYSIMQACWA 894
```



REFERENCE/DOCKET NUMBER: 990008.446C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-622-4900  
TELEFAX: 206-682-6031  
TELEX: 3723836  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1089 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 36:  
US-09-955-363-36

Query Match 22.1%; Score 1166; DB 9; Length 1089;  
Best Local Similarity 30.3%; Pred. No. 6.8e-70;  
Matches 336; Conservative 165; Mismatches 365; Indels 242; Gaps 36;  
QY 12 PLLVFSAMIFG---TITNODLPVI-----KCVLINHNKNDSSVSKSS---SYPMVSES 59  
DB 6 PAFVLGCLLTGLSLILCOLSLPILPNEKVKVQLNSFSRLRCFGESEVSWQYPMSEEE 65  
QY 60 PEDLGCALRPQSSG---TVYEAAAVEVDVSASITLQVLVDAPGNISCLWFKHSLNCQP 116  
DB 66 SSDVEIRNEENNSGLFVTVLEVSS---ASAHT-----GLYTCYNNHTQTEEN--- 110  
QY 117 HFDLQNRGV-----VSMVILKMTQAGEYLLFFIOSEATNTILFTVSIKNTLLYTL 168  
DB 111 --ELEGRIIYVDPDPVAFVPLGWT-----DYLIVE----- 141  
QY 169 RRPYFRKMNODALVCISESVPEPIVEMVLCDSQG-----ESCKE 208  
DB 142 -----DDSAIIPCRTTDPTPV--TLHNSEGVVPASVDSRQGFNGFTVGPYIC-- 189  
QY 209 ESPAVVKEKVLHELFGTDIRCARNELRECTRLTIDLNTQPTTLQFLKVGSEPL 268  
DB 190 EATVKGKKFQTFIPNVYALK-----ATSELDMEALKTV-----YKSGETI 231  
QY 269 WIRCKAVHNHGGFLAW-----ELENKA-----LEGNYFEMSTYSTNRTMIRLPFAF-VSSV 320  
DB 232 VVTC-AFVNEVVDLQWYGEVKGKGITILE-----IKVSIKLYTLTLPDPEA 280  
QY 321 ARNDGYTCSSK-----HPQSALVTIVGKGFINATNSSEDEY-IDOYEFCFSYRFK 374  
DB 281 TVKSDGYECAAQATREVKEMKVTISVHEKGFIEIKPTFSOLEAVNLHVKHFVFEVR 340  
QY 375 AYPQIRCTWTFRSKSPCEQKGLDNGYSIS-----KFCNKHQO-PGEYIP 418  
DB 341 AYPPIPRISLKNLTL-----IENLTETITDVEKIQEIRYRSKLLIRAKEEDSGHYTI 394  
QY 419 HAENDDAQFTKMTFLNRRKPOVL-----AEASASQASCFSDGYPLPSWTWKKCSDKS 471  
DB 395 VAGNEDAVKSYTFELLTQVPSLIDLVDHGGSTGGTQVRCETAGTPLPDIEWICKD-I 453  
QY 472 PNCITEITEGVWNRKANRKFVGOWSS-----STLNMSEAIKGLVKKCCAY 517  
DB 454 KCKNNETS---WTILANN-----VSNIIPIHSRORSTVEGRVTFPAKVEETIAVRCLAK 504  
QY 518 NSLGTSCETILLNSPGPPFIQDNIISFYATIGVCLLPIVVLTLICHKKYKQFPYESQLQ 577  
DB 505 NLLGAENRELKIVA---PTLRSELVAAAVLLVLIIVLIIISLIVLVVIRKQKPYEIRWR 560  
QY 578 MVQVTCSSDNEXPYVDFREYEXDLKWEFFRENLEFGKVLGSGAGFKYMMNATAYGISKTGV 637  
DB 561 VIETISPDGHEIYVDPMLPVDVSRWFEFFPKDGLVGLVGSAGFKVGEATAYGLSRSP 620  
QY 638 SIQVAYKMLKEKADSSEREAALMELKMTOLGSHENIVNLLGACTLSGGPIYLIFEYCCYG 697  
DB 621 VMKVAVKMLKPTARSEKQALMSELKIMTHLGPLNLVNLGACTKSGPIYIITEYCYFG 680  
QY 698 DLLNYLRSKREKF-----HRTWTEIFKEHNFSPYFTFQSHPN 734  
DB 681 DLVNYLHKNRDSFLSHHPEKPKKPELDIFGLNPADESTRSVILSFENNNGYMDMKQADTT 740

RESULT 11  
US-09-769-987-2  
; Sequence 2, Application US/09769987  
; Patent No. US20020055129A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsui, Toshimitsu  
; APPLICANT: Aaronson, Stuart A.  
; TITLE OF INVENTION: Antibody.  
; FILE REFERENCE: 14014. 0266U2  
; CURRENT APPLICATION NUMBER: US/09/769,987  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 08/460,656  
; PRIOR FILING DATE: 1995-06-02  
; PRIOR APPLICATION NUMBER: US 08/439,095  
; PRIOR FILING DATE: 1995-05-11  
; PRIOR APPLICATION NUMBER: US 07/915,884  
; PRIOR FILING DATE: 1992-07-20  
; PRIOR APPLICATION NUMBER: US 07/308,282  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 1089  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence; No. US20020055129A1e -  
; OTHER INFORMATION: synthetic construct  
US-09-769-987-2

Query Match 22.1%; Score 1166; DB 10; Length 1089;  
Best Local Similarity 30.3%; Pred. No. 6.8e-70;  
Matches 336; Conservative 165; Mismatches 365; Indels 242; Gaps 36;  
QY 12 PLLVFSAMIFG---TITNODLPVI-----KCVLINHNKNDSSVSKSS---SYPMVSES 59  
DB 6 PAFVLGCLLTGLSLILCOLSLPILPNEKVKVQLNSFSRLRCFGESEVSWQYPMSEEE 65  
QY 60 PEDLGCALRPQSSG---TVYEAAAVEVDVSASITLQVLVDAPGNISCLWFKHSLNCQP 116  
DB 66 SSDVEIRNEENNSGLFVTVLEVSS---ASAHT-----GLYTCYNNHTQTEEN--- 110  
QY 117 HFDLQNRGV-----VSMVILKMTQAGEYLLFFIOSEATNTILFTVSIKNTLLYTL 168  
DB 111 --ELEGRIIYVDPDPVAFVPLGWT-----DYLIVE----- 141  
QY 169 RRPYFRKMNODALVCISESVPEPIVEMVLCDSQG-----ESCKE 208  
DB 142 -----DDSAIIPCRTTDPTPV--TLHNSEGVVPASVDSRQGFNGFTVGPYIC-- 189

```
QY 209 ESPAVVKEKVLHFLGTDIRCCARNELGRECTRLFTIDLNQTPOTTLPOLFLKVGCEPL 268
Db 190 EATVKGKFKQTIPFNVAALK-----ATSELDLEMEALKTV-----YKSGETI 231
QY 269 WIRCKAVHNVHGGFLTW-----ELENKA---LEEGNYFEMSTYSTNRTMIRILFAF-VSSV 320
Db 232 AVFNEVNEVDLQWTPYGEVKGKGITMLEE-----IKVPSIKLVTLTLPVEA 280
QY 321 ARNDGYTTCSSSK-----HPSQSALVTIVGKGFINATNSSEDEYE-IDOYEFCFSVRFK 374
Db 281 TVKDSGDYECARQAOTREVEMKKTIVSHVEKGFIEIKPTFSOLEAVNLHEVHFVVEVR 340
QY 375 AYPQIRCTWTFESKSPCEQKGLDNGYSIS-----KFCNKHQO-PGEYIF 418
Db 341 AYPPPRISLKNLNLTL-----IENLTETDVEKIQEIYRSKLLIRAKEEDSGHYTI 394
QY 419 HAENDDAQTKMFTLNIRKPKQVL-----AASASQACSFSDGTPPLSWTWKKSCKS 471
Db 395 VAQNEADAVKSYTFELLTQVPSIILDLVDDHHGSGTGQVTRCTAEGTPLPIENWICKD-I 453
QY 472 PNCTEEITEGVNKRKANRKFVGOWVSS-----STLNKSEAIGFLVKCCAY 517
Db 454 KCCNNETS---WTILANN-----VSNITIEHSRDRSTVEGRVTFARVEETIAVRCLAK 504
QY 518 NSLGTSCETILLNSPGPFPIODNISFYATIGVCLLFIIVVLTLTLLCHIKYKQFYESOLQ 577
Db 505 NULLGAENRELKVA---PTLRSELVAAAVLLVLIIVLIVLVIWVKQPRYEIRWR 560
QY 578 MVQVTGSSDNEYFYVDREYEDLKWEEPPRENLEFGKVLGSGAFGKVMNATAYGISYTG 637
Db 561 VIESIPDGHIEYIYVDPNQLPYDSRWEFFPRDGLVGRVLGSGAFGKVMNATAYGISYTG 620
QY 638 SIOVAVKMLKEKADSEREALMSLKMWTOLGSHENIVNLLGACTLSGPIYLIIFYCCYG 697
Db 621 VMKAVKMLKPTARSSEKQALMSLKMTHLGLPHNLIVNLLGACTLSGPIYLIIFYCCYG 680
QY 793 YQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVICKDFGLARDIMSDSNVYVGRNARLPV 852
Db 800 YQVARGMEFLASKNCVHRDLAARNVLVLAQKIVKICDFGLARDIMSDSNVYVGRNARLPV 859
QY 853 KWMAPELSEGIYTIKSDVMSYGILLWEIFSLGVNYPGIPVDANFYKLIQNGFKMDQPF 912
Db 860 KWMAPEIFDNLTYTISDVMSYGILLWEIFSLGTPYPGMVVDSTFYKIKSGYRMAKPD 919
QY 913 YATEEYIIMQSCAFDSRRKRPFPNLTSLGCOLADAEAMYNQV---DGRVSECPH--- 967
Db 920 HATSEVYEIMVKWNSSEPEKRPSPFYHLSEIVENLLPQYKKSIEKIHLDLFLKSDHPA 979
QY 968 -----TYQNRPRSPREMDLGL 983
Db 980 MRVDSNAYIGVTYKNEEDKLKDWEGGL 1007
```

## RESULT 12

```
US-09-919-497-90
; Sequence 90, Application US/09191497
; Patent No. US20020105662A1
; GENERAL INFORMATION:
; APPLICANT: Muttler, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919, 497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
```

```
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 90
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-919-497-90
```

Query Match 22.1%; Score 1166; DB 10; Length 1089;

Best Local Similarity 30.3%; Pred No. 6.8e-70;

Matches 336; Conservative 165; Mismatches 365; Indels 242; Gaps 36;

QY 12 PLLVVFSAIFG---TITNODLPVI-----KCVLIHKNNDSSVGKSS---SYPMVSES 59

Db 6 PAFIVLGLLTLGLSLQLSLPSILPNEKVVQLNSSFSLRCFGESEVSWQPMSEE 65

QY 60 PEDLGCALRPOSSG---TVYEAAAVEVDVSAITLQVLVDAPGNISCLWFKHSLNCQP 116

Db 66 SSDVEIRNEENNSGLFVTVLEVSS---ASAAHT-----GLYTCYINHQTQTEEN--- 110

QY 117 HFDLQNRGV-----VSMVILKMTETQAGEYLLFIQSEATNTVILFTVTSIRNTLLYTL 168

Db 111 --ELEGRIIYVDPDPAFVPLGNT-----DYLVIIVB----- 141

QY 169 RRPYFRKMENODALVCISESYPEPIVEMVLCDSQG-----ESCKE 208

Db 142 -----DDDSALIPCRITTDPTPV--TLHNSGVSPASVDSRQGFNGTFTVGPYIC-- 189

QY 209 ESPAVVKEKVLHFLGTDIRCCARNELGRECTRLFTIDLNQTPOTTLPOLFLKVGCEPL 268

Db 190 EATVKGKFKQTIPFNVAALK-----ATSELDLEMEALKTV-----YKSGETI 231

QY 269 WIRCKAVHNVHGGFLTW-----ELENKA---LEEGNYFEMSTYSTNRTMIRILFAF-VSSV 320

Db 232 AVFNEVNEVDLQWTPYGEVKGKGITMLEE-----IKVPSIKLVTLTLPVEA 280

QY 321 ARNDGYTTCSSSK-----HPSQSALVTIVGKGFINATNSSEDEYE-IDOYEFCFSVRFK 374

Db 281 TVKDSGDYECARQAOTREVEMKKTIVSHVEKGFIEIKPTFSOLEAVNLHEVHFVVEVR 340

QY 375 AYPQIRCTWTFESKSPCEQKGLDNGYSIS-----KFCNKHQO-PGEYIF 418

Db 341 AYPPPRISLKNLNLTL-----IENLTETDVEKIQEIYRSKLLIRAKEEDSGHYTI 394

QY 419 HAENDDAQTKMFTLNIRKPKQVL-----AASASQACSFSDGTPPLSWTWKKSCKS 471

Db 395 VAQNEADAVKSYTFELLTQVPSIILDLVDDHHGSGTGQVTRCTAEGTPLPIENWICKD-I 453

QY 472 PNCTEEITEGVNKRKANRKFVGOWVSS-----STLNKSEAIGFLVKCCAY 517

Db 454 KCCNNETS---WTILANN-----VSNITIEHSRDRSTVEGRVTFARVEETIAVRCLAK 504

QY 518 NSLGTSCETILLNSPGPFPIODNISFYATIGVCLLFIIVVLTLTLLCHIKYKQFYESOLQ 577

Db 505 NULLGAENRELKVA---PTLRSELVAAAVLLVLIIVLIVLVIWVKQPRYEIRWR 560

QY 578 MVQVTGSSDNEYFYVDREYEDLKWEEPPRENLEFGKVLGSGAFGKVMNATAYGISYTG 637

Db 561 VIESIPDGHIEYIYVDPNQLPYDSRWEFFPRDGLVGRVLGSGAFGKVMNATAYGISYTG 620

QY 638 SIOVAVKMLKEKADSEREALMSLKMWTOLGSHENIVNLLGACTLSGPIYLIIFYCCYG 697

Db 621 VMKAVKMLKPTARSSEKQALMSLKMTHLGLPHNLIVNLLGACTLSGPIYLIIFYCCYG 680

QY 698 DLLNLYRSKREKF-----HRTWTEIFKEHNFSTYPTFQSHPN 734

Db 681 DLVNLHKNRDSFLSHHPEKPKKELDIFGLNPADESTRSVILSFENNQDYMMDKQADTT 740

QY 735 SSMPGSRVQIHPDSQOISGLHGN--SFHSEDEIEYENQKRLEEBEEDLNVLTFEDLLCFA 792

Db 741 QYVPMLEKKEVSKYSIQRSYDRPASYYKKSMILDS--VKNLLSDDNSEGLTLLDLSFT 799

QY 793 YQVAKGMEFLFKSCVHRDLAARNVLVTHGKVKVICKDFGLARDIMSDSNVYVGRNARLPV 852

```

Db 800 YQVANGMEFLASKNCVCHRDLAARNVLLAAGKIVKICDFGLARDINHDSNYVSKGSTFLPV 859
      |||||:||||| |:|||||:|||||: ||:|||||:|||||: |||||: |:|: |||
Qy 853 KWMAPESFEGITIKSDWSYGILLWEIFSLGVNYPYGPIDANFYKLIONGFKMDQPF 912
      |||||: |:|: |||||:|||||: |||||: |:|: ||| |:|: |:|: |:|:
Db 860 KWMAPESIFDNLTYTSLDWSYGILLWEIFSLGGTPYFGMWVDSFYKIKSYERMAKPD 919
      |||||: |:|: |||||:|||||: |||||: |:|: ||| |:|: |:|: |:|:
Qy 913 YATEBIYIMOSCNADFRRKRPSPNLFSTFGCQLADAEAMYNQV--DGRVSECPH--- 967
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Db 920 HATSEVYIMVKWCNSEPEKRPSPFYHLSEIVENLLPGQYKKSYEKIHLDFLKSHPAVAR 979
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Qy 968 -----TYQNRRPFSREMDLGL 983
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Db 980 MRVSDNAYIGVTKNEEDKLDWEGGL 1007
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:

RESULT 13
US-09-866-510-2
; Sequence 2, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866.510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 2
; LENGTH: 1089
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-510-2

Query Match          22.1%; Score 1166; DB 10; Length 1089;
Best Local Similarity 30.3%; Pred. No. 6.8e-70;
Matches 336; Conservative 165; Mismatches 365; Indels 242; Gaps 36;

Qy 12 PLLVFSAMIFG---TINQDLPIV-----KCVLINHKNDSSVVGKSS---SYPMVSES 59
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Db 6 PAFLVGLLGLTSLILCOLSPILPNEKVKVQLNSSFSLRCFGESEVSWQPMSEEE 65
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Qy 60 PEDLGCALRPOSSG---TVYEAAVEVDVSASITLOVLVDAPGNISLWVFKHSLNCQP 116
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Db 66 SSDVEIRNEENNSGLFVTVLEVS-----ASAAHT-----GLTYCYNHTQTEEN-- 110
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Qy 117 HFDLQNRGV-----VSMVILKMTQTQAGEYLLFIOSEATNTILFTVSIRNTLLTYL 168
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Db 111 --ELEGRIHYIVDPDPAVFPLGNT-----DYLVIIVE----- 141
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Qy 169 RPYFRKMENODALVCISESVPEPIVWVLCDSSG-----ESCKE 208
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Db 142 -----DDDSAIIPCRITTDPTPVP--TLHNSGVVPASYSRQGFNGFTVGPYIC-- 189
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Qy 209 ESPAVVKKEEKVHLFCTDIRCCARNELGRECTRLFTIDLNQTPOTLPOLFLKVGEP 268
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Db 190 EATVKGKKFQPIPNVYALK-----ATSELDLEMEALKTV-----YKSGETI 231
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Qy 269 WIRCKAVHNVHGFLTW----ELENKA---LEEENYEMSTYSTNRTMIRILFAF-VSSV 320
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Db 232 VVTC-AVFNNEVDLQWTYPYGVKGGITMLEE-----TKVSIKLUVYILTVP 280
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Qy 321 ARNDTGYYTCSKK-----HPSQSALVTIVGKGFATNNSSEDEY-IDQYEEFCFSVRFK 374
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Db 281 TVKDSGDYECAARQATREVKEMKKVTVISVHEKGFIETKPTFSQLEAVNLHEVKHFVVEVR 340
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:
Qy 375 AYPQIRCTWTSRKSFPCEQKGLONGYSIS-----KFCNKHQK--PGEYIF 418
      |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|: |:|:

```

341	Db	Y	APPPRISWLKNNITL-----IENLTETDTEVKEIQEIRYRSKLLAIRAKEEDSGHYTI 3394
419	Qy	Ha	ENDDAQTKMFTNIRRKPOVL-----ABASQASQSCFSDGYPLPSWTWKKCSDKS 471
395	Db	Y	VAQNEADVKSVTPELLTQVSSILDLVDDHHGGTGVCTVCTAGETPLPDIEHWICKD-I 453
472	Qy	Pn	CTEETIEGVVNNRKANKRVFGQWSS-----STLNSEAIKGLFLVRCCKAY 517
454	Db	Kc	NNETS---WTILANN-----VSNITIEHSRDRSTVEGRVTFKAVEETIAVRCLAK 504
518	Qy	Ns	TCSETILLNSPGPPFTIQDNISFYATIGVCLLFIVVLTLCHKYYKKQFRYSQLO 577
505	Db	Nl	GAENRELKVA-----PTLRSELTVAAAVLVLVLIISLVLVIVVIMKOKPFIIRWR 560
578	Qy	Mv	VTGSSDNEYFYVDREYEDLKWEFEPRENLEFGKVLGSGAGFKVMNATAYGISKTGV 637
561	Db	Vi	ISPDGHEYIYVDMQPLPYDSRWEPFRDGLVGLRVLGSGAGFKVVBGTATGLRSQP 620
638	Qy	Si	QAVAKMLKEKADSSREALMSCLKMNTOLGSHENIVNLLGACTLSGPIYLIFEYCCY 697
621	Db	Vm	VAKMLKPTARSESEKQALMSCLKMTIMHLGPHLNIIVNLLGACTKSGPIYIITEYCFY 680
698	Qy	Dl	NYLSKREKF-----HRTVTEIFKEHNFSFYPTFOSHNP 734
681	Db	Dl	VYVNLKHNKDSFSLSHPEKPKKELDIFGLNPADESTRSYVILSFENNGDYMDMKQADTT 740
735	Qy	Ss	MGSRVQV:HPDSDQISGLHN--SFHSEDETEYENQKRLSEEDLNVLTFEDLLCPA 792
741	Db	Qy	VPMLEKVKYSYDQISYDRPSYKKKMLDSE-VKNLLSDNSEGLTLLDLLSET 799
793	Qy	Yq	AGMEFLEFKSCVHRDLAARNVLYTHGKVKVKICDFGLARDIMSDSNVVRGNARLPV 852
800	Db	Yq	VARGMEFLASKNCVHRDLAARNVLLAQCKIYKICDFGLARDIMHDSNVVSKSTFLPV 859
853	Qy	Kw	APESLFEGIYTIKSDVWSYIGILLWEIFSLGVNVPYGPVDPANFYKLIQNGFKMDQPF 912
860	Db	Kw	APESIFDNLXYTLSDVWSYIGILLWEIFSLGCTPYPGMMVDSFTFYNKIKSGYRMAKPD 919
913	Qy	Yt	EYIIMQSCWAPDSRRKPSFNLTSFLGCOLADAEEMQNV--DGRVSECPH--- 967
920	Db	Ht	SEVEIYVWKVNSPEKRPSPFYHUSEIVENLLPGQYKKSTEKIHLDPLKSDHPAVAR 979
968	Qy	-----	TYQNRPRFSPREMDLGL 983
980	Db	MR	VDSDNAYIGVTYKNEEDKLKDWEGSL 1007
RESULT 14			
US-09-866-510-10			
; Sequence 10, Application US/09866510			
; Patent No. US20020111304A1			
; GENERAL INFORMATION:			
; APPLICANT: KAZLAUSKAS, ANDRIUS			
; APPLICANT: IKUNO, YASUSHI			
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISE			
; FILE REFERENCE: ERM-104.01			
; CURRENT APPLICATION NUMBER: US/09/866,510			
; CURRENT FILING DATE: 2001-05-25			
; PRIOR APPLICATION NUMBER: 60/250,747			
; PRIOR FILING DATE: 2000-12-01			
; PRIOR APPLICATION NUMBER: 60/289,103			
; PRIOR FILING DATE: 2001-05-07			
; NUMBER OF SEQ ID NOS: 33			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 10			
; LENGTH: 1089			
; TYPE: PR			
; ORGANISM: Homo sapiens			
US-09-866-510-10			
Query Match			22.1%; Score 1163; DB 10; Length 1089;
Best Local Similarity			30.2%; Pred. No. 1.1e-69;
Matches 335; Conservative 166; Mismatches 365; Indels 247; Gaps			

Query Match	22.1%;	Score 1163;	DB 10;	Length 1089;
Best Local Similarity	30.2%;	Pred. No. 1.1e-69;		
Matches 335;	Conservative 166;	Mismatches 365;	Indels 242;	



QY 12 PLLVFSAMIFG---TITNODLPVI-----KCVLINHKNDSSVKGSS---SYPMVSES 59  
DB 6 PAFLVGLCLTGLSLCOLSLPILPNEKENVQVQNSFSLRFCGESEVSWQPMSEEE 65  
QY 60 PEDLGCALRPOSSG---TVYEAADVVDVSATITLOVLVDAPGNISCLWVFKHSSLNCP 116  
DB 66 SSDVEIRNEENNSGLFVTVLEVSS---ASAHT-----GLYTCYNNHTQTEEN--- 110  
QY 117 HFDLQNGV-----VSMVILKMTQAGEYLLFIQSEATNYTLFTVSRINTLLYL 168  
DB 111 --ELEGRIHYIYVDPDVAFLPGLMT-----DYLIVIVE----- 141  
QY 169 RPYFERMENQDALVCISESVPEIENVLCDSOG-----ESCKE 208  
DB 142 -----DSDSALIPCRITDTPETPV--TLHNSGVVPASYSRQGFNGFTVGPYIC-- 199  
QY 209 ESPAVVVKKEEVLHELFTQDIRCCARNELGRECTRLFTIDLNTQPTTLPLQLFKVGEPL 268  
DB 190 EATVKGKKFQTIPIENVYALK---ATSELDLEMEALKTV-----YKSGETI 231  
QY 269 WIRCKAVHNVHGFGLTW-----ELEKNA---LEEGNYFEMSTYSTNRTMIRILFAF-VSSV 320  
DB 232 VVTC-AVFNNVVDLQWTPGVEVKGKITMLEE-----IKVPSIKLVYTLTVPEA 280  
QY 321 ARNDGYTTCSSK-----HPSQSALVTIVGKGFINATNSSDYE-IDOYEEFCFSVRFK 374  
DB 281 TVKDSGDYECAARQATREVKEMKKVTISVHEKGFIEIKPTFSOLEAVNLHEVHFVVEVR 340  
QY 375 AYPQIRCTWTFSRKSPCEQKGLDNGYSIS-----KPCNKHQO-PGEYIF 418  
DB 341 AYPPIRISLKNLTL-----IENLTITDVEKIQIRYRSLKLLIRAKEEDSGHYTI 394  
QY 419 HAENDDAQFTKMTLNIRKPOVL-----AEASASQACFSQDGPPLSWTWKCKSDKS 471  
DB 395 VAQNEADAVKSYTELLTQVPSSILDVDDHGGSTGGTVRCTAEGTLPDIEWMICKD-I 453  
QY 472 PNCTEEITGVWNRKANRKFVQWVSS-----STLNKSEAIGFLVKCCAY 517  
DB 454 KCKNNETS---WTILANN-----VSNITIEHSRDRSTVEGRVTFKAVETIAVRCLAK 504  
QY 518 NSLGTSCETILLNSPGPPFIQDNISFYATIGVCLLFTIWLTLILCHKYKQFRVESQLO 577

QY 12 PLLVFSAMIFG---TITNODLPVI-----KCVLINHKNDSSVKGSS---SYPMVSES 59  
DB 6 PAFLVGLCLTGLSLCOLSLPILPNEKENVQVQNSFSLRFCGESEVSWQPMSEEE 65  
QY 60 PEDLGCALRPOSSG---TVYEAADVVDVSATITLOVLVDAPGNISCLWVFKHSSLNCP 116  
DB 66 SSDVEIRNEENNSGLFVTVLEVSS---ASAHT-----GLYTCYNNHTQTEEN--- 110  
QY 117 HFDLQNGV-----VSMVILKMTQAGEYLLFIQSEATNYTLFTVSRINTLLYL 168  
DB 111 --ELEGRIHYIYVDPDVAFLPGLMT-----DYLIVIVE----- 141  
QY 169 RPYFERMENQDALVCISESVPEIENVLCDSOG-----ESCKE 208  
DB 142 -----DSDSALIPCRITDTPETPV--TLHNSGVVPASYSRQGFNGFTVGPYIC-- 199  
QY 209 ESPAVVVKKEEVLHELFTQDIRCCARNELGRECTRLFTIDLNTQPTTLPLQLFKVGEPL 268  
DB 190 EATVKGKKFQTIPIENVYALK---ATSELDLEMEALKTV-----YKSGETI 231  
QY 269 WIRCKAVHNVHGFGLTW-----ELEKNA---LEEGNYFEMSTYSTNRTMIRILFAF-VSSV 320  
DB 232 VVTC-AVFNNVVDLQWTPGVEVKGKITMLEE-----IKVPSIKLVYTLTVPEA 280  
QY 321 ARNDGYTTCSSK-----HPSQSALVTIVGKGFINATNSSDYE-IDOYEEFCFSVRFK 374  
DB 281 TVKDSGDYECAARQATREVKEMKKVTISVHEKGFIEIKPTFSOLEAVNLHEVHFVVEVR 340  
QY 375 AYPQIRCTWTFSRKSPCEQKGLDNGYSIS-----KPCNKHQO-PGEYIF 418  
DB 341 AYPPIRISLKNLTL-----IENLTITDVEKIQIRYRSLKLLIRAKEEDSGHYTI 394  
QY 419 HAENDDAQFTKMTLNIRKPOVL-----AEASASQACFSQDGPPLSWTWKCKSDKS 471  
DB 395 VAQNEADAVKSYTELLTQVPSSILDVDDHGGSTGGTVRCTAEGTLPDIEWMICKD-I 453  
QY 472 PNCTEEITGVWNRKANRKFVQWVSS-----STLNKSEAIGFLVKCCAY 517  
DB 454 KCKNNETS---WTILANN-----VSNITIEHSRDRSTVEGRVTFKAVETIAVRCLAK 504  
QY 518 NSLGTSCETILLNSPGPPFIQDNISFYATIGVCLLFTIWLTLILCHKYKQFRVESQLO 577

QY 968 -----TYQNRPRPFSREMDLGL 983  
DB 980 MRYVDSNAVIGVTYKNEEDKLDWEGGL 1007

RESULT 15  
US-09-866-510-4  
; Sequence 4, Application US/09866510  
; Patent No. US20020111304A1  
; GENERAL INFORMATION:  
; APPLICANT: KAZLAUSKAS, ANDRIUS  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES  
; FILE REFERENCE: ERM-104.01  
; CURRENT APPLICATION NUMBER: US/09/866,510  
; CURRENT FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 60/250,747  
; PRIOR FILING DATE: 2000-12-01  
; PRIOR APPLICATION NUMBER: 60/289,103  
; PRIOR FILING DATE: 2001-05-07  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 1089  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-866-510-4

Query Match 22.0%; Score 1162; DB 10; Length 1089;  
Best Local Similarity 30.2%; Pred. No. 1.3e-69;  
Matches 335; Conservative 166; Mismatches 365; Indels 242; Gaps 36;

QY 12 PLLVFSAMIFG---TITNODLPVI-----KCVLINHKNDSSVKGSS---SYPMVSES 59  
DB 6 PAFLVGLCLTGLSLCOLSLPILPNEKENVQVQNSFSLRFCGESEVSWQPMSEEE 65  
QY 60 PEDLGCALRPOSSG---TVYEAADVVDVSATITLOVLVDAPGNISCLWVFKHSSLNCP 116  
DB 66 SSDVEIRNEENNSGLFVTVLEVSS---ASAHT-----GLYTCYNNHTQTEEN--- 110  
QY 117 HFDLQNGV-----VSMVILKMTQAGEYLLFIQSEATNYTLFTVSRINTLLYL 168  
DB 111 --ELEGRIHYIYVDPDVAFLPGLMT-----DYLIVIVE----- 141  
QY 169 RPYFERMENQDALVCISESVPEIENVLCDSOG-----ESCKE 208  
DB 142 -----DSDSALIPCRITDTPETPV--TLHNSGVVPASYSRQGFNGFTVGPYIC-- 189  
QY 209 ESPAVVVKKEEVLHELFTQDIRCCARNELGRECTRLFTIDLNTQPTTLPLQLFKVGEPL 268  
DB 190 EATVKGKKFQTIPIENVYALK---ATSELDLEMEALKTV-----YKSGETI 231  
QY 269 WIRCKAVHNVHGFGLTW-----ELEKNA---LEEGNYFEMSTYSTNRTMIRILFAF-VSSV 320  
DB 232 VVTC-AVFNNVVDLQWTPGVEVKGKITMLEE-----IKVPSIKLVYTLTVPEA 280  
QY 321 ARNDGYTTCSSK-----HPSQSALVTIVGKGFINATNSSDYE-IDOYEEFCFSVRFK 374  
DB 281 TVKDSGDYECAARQATREVKEMKKVTISVHEKGFIEIKPTFSOLEAVNLHEVHFVVEVR 340  
QY 375 AYPQIRCTWTFSRKSPCEQKGLDNGYSIS-----KPCNKHQO-PGEYIF 418  
DB 341 AYPPIRISLKNLTL-----IENLTITDVEKIQIRYRSLKLLIRAKEEDSGHYTI 394  
QY 419 HAENDDAQFTKMTLNIRKPOVL-----AEASASQACFSQDGPPLSWTWKCKSDKS 471  
DB 395 VAQNEADAVKSYTELLTQVPSSILDVDDHGGSTGGTVRCTAEGTLPDIEWMICKD-I 453  
QY 472 PNCTEEITGVWNRKANRKFVQWVSS-----STLNKSEAIGFLVKCCAY 517  
DB 454 KCKNNETS---WTILANN-----VSNITIEHSRDRSTVEGRVTFKAVETIAVRCLAK 504  
QY 518 NSLGTSCETILLNSPGPPFIQDNISFYATIGVCLLFTIWLTLILCHKYKQFRVESQLO 577

```

  |||
505 NLGAENRELKVA-----PTRLSELVAAVLVLLVIVISLIVVIMVKQKPYEIRWR 560
QY 578 MVQVTCSSNEYFYVDFREYEDLKWEFFRENLEFGKVLGSGAFGKVMNATAYGISTGV 637
Db 561 VIESISPDGHEIYVDPMLPYDSRWKFPDRDGLVGLGSGAFGKVEGTAYGLSRSP 620
QY 638 SIQVAVKMLKEKADSSERELMSLKMOTOLGSHENIVNLGACTLSGPIYLIFEYCCYG 697
Db 621 VMKVAVKMLKPTARSEKQALMSLKMIMHLGPHLNIYNLLGACTKSGPIIITEYCFYG 680
QY 698 DLNLYLRSRREKF-----HRTWTEIFKEHNFSEFYPTFQSHPN 734
Db 681 DLVNYLHKNRDSFLSHHPEKPKKELDIFGLNPADESTRSVILSEENNGYMDMKQADTT 740
QY 735 SSMPCGREVQIHPDSDQISGLHGN--SFHSEDEIEYENQKRLEEBEDLVLTFTEDLLCFA 792
Db 741 QYVPMLEKEVSKYSIQIRSLYDRPASYYKKSMLDSE-VKNLLSDDNSEGLTLLDLST 799
QY 793 YQVAKGMEFLEFKSCVHRDLAARNVLVTHGKVYKICDFGLARDIMSDSNVVRGNARLPV 852
Db 800 YQVARGMEFLASKNCVHRDLAARNVLLAOGKIVKICDFGLARDIMHDSNIVSKGSTFLPV 859
QY 853 KWMAPELPEGIYTIKSDVMSYGILLWEIFSLGVNYPYGPVDPANFYKLIQNGFKMDQPF 912
Db 860 KWMAPEIFDNLVYTTLSDVMSYGILLWEIFSLGGTYPYGMWVDSTFYNTKSGYRMAKPD 919
QY 913 YATEEYIIMQSCWAFDSKRPSFPNLTSLGCOLADAEAMTONV--DGRVSECPH--- 967
Db 920 HATSEYIEIMVRCWNEPEKRPSPFYLSEIVENLLPQYKKSYEKIHLDFLKSHPAVAR 979
QY 968 -----TYQNRPFPSREMDLGL 983
Db 980 MRVSDSNAYIGVYTKNEEDKLDWEGGL 1007
```

Search completed: May 27, 2003, 14:45:10  
Job time : 29.5123 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:27:39 ; Search time 26.5134 seconds  
(without alignments)  
3600.506 Million cell updates/sec

Title: US-09-919-408-4

Perfect score: 5274

Sequence: 1 MPALARDAGTVPLLVFESAM.....PFSREMDLGLSPQAQVEDS 993

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5238	99.3	993	2 A36873	protein-tyrosine k
2	4533.5	86.0	1000	2 A38827	Flt3 protein - mou
3	4429.5	84.0	992	2 A39931	protein-tyrosine k
4	1286	24.4	977	2 I45877	protein-tyrosine k
5	1266	24.0	975	1 TVMSKT	protein-tyrosine k
6	1263.5	24.0	978	1 A49814	protein-tyrosine k
7	1251	23.7	976	1 TVHUKT	protein-tyrosine k
8	1226	23.2	954	2 I51703	c-kit-related kina
9	1225.5	23.2	980	1 TVCTMD	macrophage colony-
10	1218	23.1	941	1 TVHVMMD	protein-tyrosine k
11	1216.5	23.1	972	1 TVHUMD	macrophage colony-
12	1215	23.0	960	1 JN0677	protein-tyrosine k
13	1213	23.0	975	2 T30816	macrophage colony-
14	1194	22.6	978	2 S16385	macrophage colony-
15	1183.5	22.4	976	1 TVMSMD	macrophage colony-
16	1176	22.3	1088	1 PFRTGA	platelet-derived g
17	1166	22.1	1089	1 PFHUGA	platelet-derived g
18	1146	21.7	1089	1 S33727	platelet-derived g
19	1128	21.4	1087	2 I51552	platelet-derived g
20	1105.5	21.0	1098	1 PMSRB	platelet-derived g
21	1086.5	20.6	1106	1 PFHUGB	platelet-derived g
22	1080	20.5	1048	2 T30815	platelet-derived g
23	992	18.8	790	1 F0MVHZ	gag-kit polyprotei
24	990	18.8	1338	2 S09982	protein-tyrosine k
25	959.5	18.2	1333	2 I76875	receptor tyrosine
26	956	18.1	1336	2 I60598	Fit-1 tyrosine kin
27	944.5	17.9	1356	2 J61402	protein-tyrosine k
28	939.5	17.8	1330	2 S49010	embryonic receptor
29	935	17.7	1348	2 S51656	vascular endotheli

30	932	17.7	1379	2 JC4954	vascular endotheli
31	923	17.5	1367	2 A41228	protein-tyrosine k
32	907.5	17.2	1363	2 I58375	protein-tyrosine k
33	904.5	17.2	1298	2 A48999	protein-tyrosine k
34	874.5	16.6	823	2 B35963	protein-tyrosine k
35	851	16.1	160	2 A39061	protein-tyrosine k
36	849	16.1	821	1 TVHUF2	fibroblast growth
37	842	16.0	813	1 A49123	fibroblast growth
38	841	15.9	824	2 S24108	protein-tyrosine k
39	836.5	15.9	822	2 A45081	fibroblast growth
40	832.5	15.8	824	2 S36439	fibroblast growth
41	828.5	15.7	769	2 S16236	fibroblast growth
42	828.5	15.7	822	2 A41794	keratinocyte growth
43	827	15.7	797	2 S38579	fibroblast growth
44	825	15.6	821	1 TVMSBK	fibroblast growth
45	823.5	15.6	822	2 B54846	fibroblast growth

ALIGNMENTS

RESULT 1

A36873  
protein-tyrosine kinase (EC 2.7.1.112), STK-1 precursor - human  
N:Alternate names: stem cell tyrosine kinase 1  
C:Species: Homo sapiens (man)  
C:Date: 03-Nov-1995 #sequence\_revision 03-Nov-1995 #text\_change 04-Feb-2000  
C:Accession: A36873  
R:Small, D.; Levenstein, M.; Kim, E.; Carow, C.; Amin, S.; Rockwell, P.; Witte, L.; B  
Proc. Natl. Acad. Sci. U.S.A. 91, 459-463, 1994  
A:Title: STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in CD34(+)  
A:Reference number: A36873; MUID:94119906; PMID:7507245  
A:Accession: A36873  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-993 <SMA>  
A:Cross-references: GB:U02687  
A:Note: In the authors translation, an additional residue Ala is shown after 420-Ala  
C:Genetics:  
A:Map position: 13q12  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom  
C:Keywords: ATP; glycoprotein; phosphotransferase; transmembrane protein; tyrosine-sp  
F:608-950/Domain: protein kinase homology <KIN>  
F:616-624/Region: protein kinase ATP-binding motif

Query Match	99.3%	Score	5238	DB	2	Length	993
Best Local Similarity	99.7%	Pred. No.	2.3e-247				
Matches	991	Conservative	0	Mismatches	1	Indels	2
						Gaps	2
QY	1	MPALARDAGTVPLLVFESAMIFGTITNQDLPVTKCVLINHKNDSSVKGSSSPVMSSESP	60				
Db	1	MPALARDAGTVPLLVFESAMIFGTITNQDLPVTKCVLINHKNDSSVKGSSSPVMSSESP	60				
QY	61	EDLGCALRPOSSCTVVEAAAVEVDVSASITLQVLVDAPGNISCLWFKVHSLNCQPHFDL	120				
Db	61	EDLGCALRPOSSCTVVEAAAVEVDVSASITLQVLVDAPGNISCLWFKVHSLNCQPHFDL	120				
QY	121	ONRGVVSWMILKMTETQAGEYLLFIQSEATNYTLFTVSIRNTLLYTLRRPYFRKMNQD	180				
Db	121	ONRGVVSWMILKMTETQAGEYLLFIQSEATNYTLFTVSIRNTLLYTLRRPYFRKMNQD	180				
QY	181	ALVCISESPPEPVEVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE	240				
Db	181	ALVCISESPPEPVEVWVLCDSQGESCKEESPAVVKKEKVLHFLGTDIRCCARNELGRE	240				
QY	241	CTRLFTIDLNOTPQTTLPOLFLKVGEPWLWRCRAVHVHNGFGLTWELNKALEGNCVFEM	300				
Db	241	CTRLFTIDLNOTPQTTLPOLFLKVGEPWLWRCRAVHVHNGFGLTWELNKALEGNCVFEM	300				
QY	301	STYSTNRTMIRILFAFVSSVARNDTGYYTCSSSKHPQSOSALVTIVGKGFINATNSSDYE	360				
Db	301	STYSTNRTMIRILFAFVSSVARNDTGYYTCSSSKHPQSOSALVTIVGKGFINATNSSDYE	360				



A:Reference number: A39931; MUID:91292518; PMID:1648448  
A:Accession: A39931  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-992 <MAT>  
C:Cross-references: GB:M64689; NID:g193327; PIDN:AAA37634.1; PID:g193328  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo  
C:Keywords: ATP; phosphotransferase; transmembrane protein; tyrosine-specific protein ki  
F:609-953/Domain: protein kinase homology <KIN>  
F:617-625/Region: protein kinase ATP-binding motif

Query Match 84.0%; Score 4429.5; DB 2; Length 992;  
Best Local Similarity 84.1%; Pred. No. 4.6e-208;  
Matches 836; Conservative 57; Mismatches 90; Indels 11; Gaps 4;  
QY 1 MPALA-RDAGTVPLLVFSAMIFGTTNODLPVIRKVLINHNKNDSSVGKSSSPVMSVES 59  
DB 1 MRALQSRDRLLLVLSMILETVTQDLVIRKVLISHENSSACKPSRYMRGS 60  
QY 60 PEDIGCALRPOSSGTVYEAAYVDDVSASITLOVLVDAPGNISCLVFKHSLNCPHFED 119  
DB 61 PEDLOCTPRROSEGVYEAATVAESGSILOVQLATPGDLSCLVWFKHSLGCGPHFD 120  
QY 120 LQNRGVSMVILKMTQAGEYLLFIQSEATNTILFTVIRNTLLYLRRPYFRKMEQ 179  
DB 121 LQNRGVSMVILKMTQAGEYLLFIQSEATNTILFTVIRNTLLYLRRPYFRKMEQ 180  
QY 180 DALVCISEVPEPIVWVLCDSQGSKEESPVAVKKEKVLHEFGTDIRCARNELGR 239  
DB 181 DALLCISEGVPEPTVWVLCSSHRESKEEGPAVVRKKEKVLHEFGTDIRCARNALGR 240  
QY 240 ECTRLFTDLNQTPOTTPOLPLKVGEPDLWIRKAVHVNHGGLTWELNKALEEGNYFE 299  
DB 241 ECTRLFTDLNQTPOTTPOLPLKVGEPDLWIRKAVHVNHGGLTWELNKALEEGNYFE 300  
QY 300 MSTYSTNTMIRILAFVSSVARNDTGYTCSSSRHPSQSALVTIVGKGFINATSSDEY 359  
DB 301 MSTYSTNTMIRILAFVSSVARNDTGYTCSSSRHPSQSALVTIVGKGFINATSSDEY 360  
QY 360 EIDQYEEFCFVRFKAYPOIRCTTFRSRKSPCEQKGLDNGYSISKFCNKKHQPGEYIF 419  
DB 361 EIDPYEKCFVRFKAYPRICTWIFSQASPPCEQGLEGGYSISKFCNKKHQPGEYIF 420  
QY 420 AENDDAQTKMFTLNIRKPKQVLASASQASCFSDGYPPLSWTKKCSKDPNCTEET 479  
DB 421 AENDDAQTKMFTLNIRKPKQVLASASQASCFSDGYPPLSWTKKCSKDPNCTEET 480  
QY 480 EGVNKRANRVFGQVSSSTLNMSSEATKGLVKCCAYNSLCTSETILLNSPGPPFIQ 539  
DB 481 EGVNKRANRVFGQVSSSTLNMSSEATKGLVKCCAYNSLCTSETILLNSPGPPFIQ 540  
QY 540 DNISFYATIGVCLLPFIVVLLIICHYKQPRYESQLQMVQVGTGSSDNEYFYVDFREY 599  
DB 541 DNISFYATIGVCLLPFIVVLLIICHYKQPRYESQLQMVQVGTGSSDNEYFYVDFREY 600  
QY 600 DLKWEFPRENLEFGKVLGSGAPCKVYNATAYIGTKGVSIQVAVKMLKADSSREALM 659  
DB 601 DLKWEFPRENLEFGKVLGSGAPCKVYNATAYIGTKGVSIQVAVKMLKADSSREALM 660  
QY 660 SELKMTOLGSHENTVNLGACTLSGPVILFEYCCYGDLLNVLRSKRFHRTWTEIFK 719  
DB 661 SELKMTOLGSHENTVNLGACTLSGPVILFEYCCYGDLLNVLRSKRFHRTWTEIFK 720  
QY 720 EHNFSFYTFQSHPNSSMPGSRVQIHDPDQISGLHNSFSEDEIEYENOKRL--EEE 777  
DB 721 EHNFSFYTFQSHPNSSMPGSRVQIHDPDQISGLHNSFSEDEIEYENOKRLAE 780  
QY 778 EDNLVLTEDLLCFAYQAKMEFLFKSCVHRDLAARNVLVTHGKVKVLCDFGLARDIM 837  
DB 781 EDNLVLTEDLLCFAYQAKMEFLFKSCVHRDLAARNVLVTHGKVKVLCDFGLARDIL 840  
QY 838 SDSNVYVRGNARLPVKWAPESLFGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 897  
DB 839 SDSNVYVRGNARLPVKWAPESLFGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 900

DB 841 SDSNVYVRGNARLPVKWAPESLFGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDAN 900  
QY 898 FYKLIONGKMDQPFYATBEIYIIMOSWAFDGRKRPSPNLTSLFGCQLADAEAMQYN 957  
DB 901 FYKLIONGKMDQPFYATBEIYIIMOSWAFDGRKRPSPNLTSLFGCQLADAEAMQYN 957  
QY 958 VDRGVSECTPTVYQNRPPSPREMDLGLLSPOAQVE 991  
DB 958 ---IRTSIHLPKQAAPQQRG-GLRAQSPORQVK 986  
RESULT 4  
I45877  
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - aurochs  
C:Species: Bos primigenius (aurochs)  
C:Date: 19-Dec-1997 #sequence\_revision 19-Dec-1997 #text\_change 20-Apr-2000  
C:Accession: I45877  
R:Kubota, T.; Hiko, H.; Sasaki, E.; Sakurai, M.  
Gene 141, 305-306, 1994  
A:Title: Sequence of a bovine c-kit proto-oncogene cDNA.  
A:Reference number: I45877; MUID:94215924; PMID:7512939  
A:Accession: I45877  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-977 <KUB>  
C:Cross-references: GB:D16680; NID:g516659; PIDN:BAA04084.1; PID:g516660  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo  
C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase  
F:329-395/Domain: immunoglobulin homology <IMM>  
F:588-932/Domain: protein kinase homology <KIN>

Query Match 24.4%; Score 1286; DB 2; Length 977;  
Best Local Similarity 33.2%; Pred. No. 2.7e-55;  
Matches 332; Conservative 172; Mismatches 344; Indels 152; Gaps 36;  
QY 47 VKGSSSPVMSVESPEDLGC-ALRPOSSGTVYEAAYVDDVSASITLOVLVDAPGNISCLW 105  
DB 20 VQSGSQSV--SPGELSLSIHPAKS-----ELIVSGVDEIRLLCTDPGFVK-W 66  
QY 106 VFKHSLNCPHFEDLQNRGVSMVILKMTQAGEYLLFIQSEAT---NTILFTVSI 162  
DB 67 TFE-----ILQLSKTNPEWIT-EKATNTGNTCTNKGGLSS 105  
QY 163 TL-----LYTLRRPYFRKMEQDALVCISEVPEPIVWVLCDSQGSKEESP 213  
DB 106 SIYVVRDPPEKFLIDLPLYGKEEN-DTLVRCPLTQPE-VTNYSLTGCEGKPLPKDLTFV 163  
QY 214 -----VKKEKVLHEFGTDIRC--CARNELGRE-CTRLETFID---LNQTPQTL 258  
DB 164 ADPKAGITIRNVKREYHRL-----CLHCSANQRKSMKSKFTLVRAAIKAVPVVSVS 217  
QY 259 Q--LFLKVGEPDLWIRKAVHVNHGGLTWELNKALEEGNYFEMSTYSTNRTMIRILFA 316  
DB 218 KTSYLLREGEFAVTCILINDVSSVSDSMWIKENSQQTAKQTKNSWHQGDYSYLQERLT 277  
QY 317 VSSVARNDTGYTCSSSKHPQSALVT---IVGKGFNA-----TNSSDEYEDIQY 364  
DB 278 ISSARNVDGVFCYANNTFGSANVTTLLEVVDKGFINIPPMNTTVFVNDGENVDL--- 334  
QY 365 EEFECFVRFKAYPO-IRCTWTFRKS-----FPCQKGLDNGYSISKFCNKKH-- 411  
DB 335 -----VVEEYAKPKPVHROWIYNNRTSTDKWDDYPKSE-----NESNIRVYNELHLTRLK 384  
QY 412 --OPGEYIFHAENDDAQTKMFTLNIRKPKQVLASASA--SQASCFSDGYPPLSWTKK 467  
DB 385 GTGGGTTHVNSVDVNSVTFVNVYNTKPEILTHDLVNGMLQCVAAGPEPTIDHYFC 444  
QY 468 SKSPNCTEET-TEGVNKRANRVFGQVSSSTLNMSSEATKGLVKVLCDFGLARDIM 526  
DB 445 PGTEQRCSVPVGVQVQIQNSSVSPGKLVVYSTIDSTEFKHNGTVECRAYNDVGKSSAS 504  
QY 527 ILLNSGPPPP-----FIQDNISFYATIGVCLLPFIVVLLIICHYKQPRYESQLQMV 578  
DB 528 ILLNSGPPPP-----FIQDNISFYATIGVCLLPFIVVLLIICHYKQPRYESQLQMV 579

Db 505 FNFAKSGSKQIHAHTLFTPLLIGFVIAAGLMCFVMTL-----YKYLQKPMYEVQWKV 560  
Qy 579 V-OVTGSSDSEYFYDFEYEDLKWEPRNLEFGKVLGSGAFKGVNATAYGSKTGV 637  
Db 561 VEEING-----NNYVYIDPTQDPYDHWKPEPRNLSFGKTLGAGAFKGVNATAYGLIKSDA 617  
Qy 638 SIQVAVKMLKEKADSSEREALMSLKMNTLGSHEINYNLGLACPLSGPIYLIFYCYCG 697  
Db 618 AMTAVKMLKPSAHLTEREALMSLKLVSILGNHNNIYNLLGACITIGPTLVITEYCYCG 677  
Qy 698 DLLNLYSKRKEKHFRTWTE-----IFKEHNFSPFTTQSHPNSSM---PGSREVOIHPD 748  
Db 678 DLLNFLRKRDSTFCSDKQEDAEVALYKLNLLHSKSSCNDSTNEYMDKMPGVSYV-VPTK 736  
Qy 749 SDQISLGHNSFHSDETEYENOKLEFEEDLVNLTFFDLGCFAYQVAKGMEFLFKSOV 808  
Db 737 ADKRRSARIGS-----IERVTPAIMEDEL-ALDLELDSFQVAKGMFLASKNCI 790  
Qy 809 HRDLAARNLVTHGKVVKICDFGLARDIMSDSNVYVRGNARLPVKWMAPESEFEGYTIK 868  
Db 791 HRDLAARNLLTHGRITKICDFGLARDIKDSDSNVYVRGNARLPVKWMAPESEFNCVTFE 850  
Qy 869 SDVWSYGILLWEIFSLGNPNPVGIPVDANFYKLIONGFKMDQPFYATEBIYIIMQSCWAF 928  
Db 851 SDVWSYGIFLWELFSLGSSPYGMPVDSKFKYKMKIEGFRMLSPHAPAEYMDIMKTCWDA 910  
Qy 929 DSRKRPSPNLTSLGCLADAEAEAMYNQVGRVSEC-PH 967  
Db 911 DPLKRPTFKQIVOLIEKOISESTNHIYSN-----LANCSPH 946

## RESULT 5

TMWSKT  
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - mouse  
N:Alternate names: tyrosine kinase receptor c-kit  
C:Species: Mus musculus (house mouse)  
C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 31-Mar-2000  
C:Accession: S00474; B44876; I49596  
R:Qiu, F.; Ray, P.; Brown, K.; Barker, P.E.; Jhanwar, S.; Ruddle, F.H.; Besmer, P.  
EMBO J. 7, 1003-1011, 1988  
A:Title: Primary structure of c-kit: relationship with the CSF-1/PDGF receptor kinase fa  
A:Reference number: S00474; MUID:88296403; PMID:2456920  
A:Accession: S00474  
A:Molecule type: mRNA  
A:Residues: 1-975 <QID>  
A:Cross-references: GB:Y00864; NID:g50423; PIDN:CAA68772.1; PID:g50424  
R:Rossi, P.; Marzali, G.; Albanesi, C.; Charlesworth, A.; Geremia, R.; Sorrentino, V.  
Dev. Biol. 152, 203-207, 1992  
A:Title: A novel c-kit transcript, potentially encoding a truncated receptor, originates  
A:Reference number: A44876; MUID:92331813; PMID:1378413  
A:Accession: B44876  
A:Molecule type: DNA  
A:Residues: 771-814 <ROS>  
A:Note: sequence extracted from NCBI backbone (NCBIN:108837, NCBI:108840)  
R:Yasuda, H.; Galli, S.J.; Geissler, E.N.  
Biochem. Biophys. Res. Commun. 191, 893-901, 1993  
A:Title: Cloning and functional analysis of the mouse c-kit promoter.  
A:Reference number: I49596; MUID:93221533; PMID:7682073  
A:Accession: I49596  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-22 <RES>  
A:Cross-references: GB:L11358; NID:g293325; PIDN:AAA37420.1; PID:g293326  
C:Genetics:  
A:Gene: kit; c-kit  
A:Map position: 5  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;  
fic protein kinase  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-975/Product: protein-tyrosine kinase kit #status predicted <KTC>  
F:23-519/Domain: extracellular #status predicted <EXT>  
F:51-100/Domain: immunoglobulin homology <IMM1>

F:130-189/Domain: immunoglobulin homology <IMM2>  
F:227-295/Domain: immunoglobulin homology <IMM3>  
F:321-397/Domain: immunoglobulin homology <IMM4>  
F:426-496/Domain: immunoglobulin homology <IMM5>  
F:520-542/Domain: transmembrane #status predicted <IMM>  
F:543-975/Domain: intracellular #status predicted <INT>  
F:586-929/Domain: protein kinase homology <KIN>  
F:594-602/Region: protein kinase ATP-binding motif  
F:58-98,137-187,234-293,431-494/Disulfide bonds: #status predicted  
F:146,296,303,323,355,370,466,489/Binding site: carbohydrate (Asn) (covalent) #status  
F:622,639,790/Active site: Lys, Glu, Asp #status predicted  
F:795,808/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 24.0%; Score 1266; DB 1; Length 975;

Best Local Similarity 33.2%; Pred. No. 2.5e-54;

Matches 336; Conservative 173; Mismatches 330; Indels 172; Gaps 41;

Qy 35 CVLINHKNDSVGVSSSPYPMVSESPEDIGCALRPOSSGTVEAAAVEVDVSITLOVL 94  
Db 12 CVLLVLLRGQTATSPSPASGPPSP-----STHPAQSS-----ELIVEAGDTL--- 54  
Qy 95 VDAPGNISCL-----WVFKHSLNCPHFEDLQNGVSMVLKMTETQAGEYLLFIQSE 148  
Db 55 -----SLTCIDPDFVKTFFK-----TYFN-----EMVENKKNEWIO-EKAE 89  
Qy 149 ATNYTLFTVTSIRNTL-----LYTLRRPYFRKMENODALVCISESVPEPIVE 195  
Db 90 ATR-TGYTCSNSNGLTSSLYVVRDPAKFLVGLPLFGK-EDSDALVRCPLTDPQ-VSN 146  
Qy 196 WVLCDSQGSCKEESPAY-----VKKEEYVHLHELFGDTRCCARNELGRCTRL--- 244  
Db 147 YSLIEDGKSLPTDLTFVPNKRAGITIKNVKRAYHRLC-----VRCAAQ-----RDGTWLHSD 199  
Qy 245 -FTIDLNQI-----PQTLPLQL--FLAVGEPLWTRCAVHVNHGFLTW-----E 286  
Db 200 KFTLVKREAIKAIPIVYVSPVETSHLLKGGDTFTVCTIKDYSTSVNSMWMKMNPOQIHQA 259  
Qy 287 LENKALEEGNPEMSTYSTNRTMIRILFAFVSSVARNDGTGYTCSSSKHPQSALVT--- 343  
Db 260 VKHNSHWRGDF-----NYRQETILT-----ISSARVDDSGVFMCIYANNTEGSAVTTLK 309  
Qy 344 IVKGFINATN-SSDEYEDQYEEFCFSVRFAKAYPO-IRCTWTFSTRKSPCEQKGLD--- 398  
Db 310 VVEKGFINISPVKNNTTFVTDGENVDLVVEYEAYPEHQOQWIYMNRT--SANKGKDYK 367  
Qy 399 -NGYSISKFCNKH-----QGEVIFHAENDDAQFTKMTLNIRKPKQVLA--EASAS 448  
Db 368 SDNKSIRIYVNLRLTLKGTGGTYTFLVNSDSASVTFNVYVNTKPEILTYDRLING 427  
Qy 449 QASCFSDGYPLPSMTWKKCKSDKSPNCTEITE-GVNNRKANRVFGQWSSSTLNMSEAI 507  
Db 428 MLOCVAGFPETIDWYFCUGAEQRCTTPSPVDVQVQNVSVSPFGKLVVQSSIDSVFR 487  
Qy 508 KGLVKKCAYNLSGTSCETILLNSPGFPFP---IQDN-----ISFYATIGVCLLFIV 557  
Db 488 HNGTVECKASNDVGKS--SAFFN---FAFKEIQLOAHTLFTPLLIGFVVAAGAMGIIVM 541  
Qy 558 LTLILCHYKKQFRESQLOMV-QVTGSSDNEYFYVDREYEDLKWFEFRENLEFKVL 616  
Db 542 LT-----YKYLQKPMYEVQWKVVEEING---NNYVYIDPTQDPYDHWKPEPRNLSFGKTL 594  
Qy 617 GSGAFKVMNATAYGTSKGVSTQVAVKMLKEKADSSEREALMSLKMNTLQSGSHENIVN 676  
Db 595 GAGAFKGVNATAYGLIKSDAAMTVAVKMLKPSAHLTEREALMSLKLVSILGNHNNIYN 654  
Qy 677 LLGACTLSGPIYLIFYCYCCYCGDLLNLSRKRKF-----HRTWTEIFKEHNFSPFTFQ 730  
Db 655 LLGACTVGGPTLVITEYCCYCGDLLNLSRKRKDSFIFSKQEQAEEAALYKLLHSTESCD 714  
Qy 731 SHPNSSM---PGSREVOIHPDSDOISGLHNSFHSDETEYENOKLEFEEDLVNLTFFDL 787  
Db 715 S-SNEYMDKMPGVSYV-VPTKDKRRSARIDSY-----IERDVTPTAIMEDEL-ALDLD 766





A:Cross-references: GB:S78839; NID:g244084; PIDN:AAB21234.1; PID:g244085  
A:Note: sequence extracted from NCBI backbone (NCBIN:78839, NCBI:P:78842)  
A:Note: disease-related mutant from patient with piebaldism  
A:Accession: B41815  
A:Molecule type: DNA  
A:Residues: 637-641; 'SPELPW' <SP2>  
A:Cross-references: GB:S78843; NID:g244086; PIDN:AAB21235.1; PID:g244087  
A:Note: sequence extracted from NCBI backbone (NCBIN:78843, NCBI:P:78844)  
A:Note: disease-related mutant from patient with piebaldism  
A:Accession: C41815  
A:Molecule type: DNA  
A:Residues: 556-560; 'GGDKWK' <SP3>  
A:Cross-references: GB:S78845; NID:g244088; PIDN:AAB21236.1; PID:g244089  
A:Note: sequence extracted from NCBI backbone (NCBIN:78845, NCBI:P:78846)  
A:Note: disease-related mutant from patient with piebaldism  
A:Giebel, L.B.; Strunk, K.M.; Holmes, S.A.; Spritz, R.A.  
Oncogene 7, 2207-2217, 1992  
A:Title: Organization and nucleotide sequence of the human KIT (mast/stem cell growth factor receptor) gene  
A:Reference number: 137948; MUID:93064697; PMID:1279499  
A:Accession: 137948  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-976 <RES>  
A:Cross-references: EMBL:X69301; NID:g34089; PIDN:CAA49159.1; PID:g825686  
A:Note: an alternative splice form omitting residues 510-513 is described  
R:Yamamoto, K.; Tojo, A.; Aoki, N.; Shibuya, M.  
Jpn. J. Cancer Res. 84, 1136-1144, 1993  
A:Title: Characterization of the promoter region of the human c-kit proto-oncogene.  
A:Reference number: 156954; MUID:94103107; PMID:7506248  
A:Accession: 156954  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-22 <RES>  
A:Cross-references: GB:S67773; NID:g450358; PIDN:AAB29529.1; PID:g450359  
R:Spritz, R.A.; Holmes, S.A.; Berg, S.Z.; Nordlund, J.J.; Fukui, K.  
Hum. Mol. Genet. 2, 1499-1500, 1993  
A:Title: A recurrent deletion in the KIT (mast/stem cell growth factor receptor) proto-oncogene  
A:Reference number: 154336; MUID:94061059; PMID:7694728  
A:Accession: 154336  
A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 242-250 <RES>  
A:Cross-references: GB:S67686; NID:g460545; PIDN:AAD13996.1; PID:g4261696  
C:Genetics:  
A:Gene: GDB:KIT  
A:Cross-references: GDB:120117; OMIM:164920  
A:Map position: 4q12-q13  
A:Introns: 23/1; 113/1; 207/1; 252/3; 309/1; 372/2; 411/1; 449/2; 514/1; 549/3; 592/1; 600/1  
A:Note: defects in this gene may result in piebaldism  
C:Function:  
A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology  
C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; kinase-related  
A:Title: tyrosine-specific protein kinase  
F:1-976/Product: protein-tyrosine kinase  
F:1-509, 514-976/Product: protein-tyrosine kinase kit precursor, long form #status predicted <NAT>  
F:23-976/Product: protein-tyrosine kinase kit precursor, short form #status predicted <NAT>  
F:23-520/Product: protein-tyrosine kinase kit #status predicted <EXT>  
F:51-99/Domain: immunoglobulin homology <IMM1>  
F:129-188/Domain: immunoglobulin homology <IMM2>  
F:226-292/Domain: immunoglobulin homology <IMM3>  
F:328-394/Domain: immunoglobulin homology <IMM4>  
F:423-493/Domain: immunoglobulin homology <IMM5>  
F:521-543/Domain: transmembrane #status predicted <TM>  
F:544-976/Domain: intracellular #status predicted <INT>  
F:587-931/Domain: protein kinase homology <KIN>  
F:595-603/Region: protein kinase ATP-binding motif  
F:58-97, 136-186, 233-290, 428-491/disulfide bonds: #status predicted  
F:130, 145, 283, 300, 320, 352, 367, 463, 486/Binding site: carbohydrate (Asn) (covalent) #s  
F:623, 640, 792/Active site: Lys, Glu, Asp #status predicted  
F:797, 810/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 23.7%; Score 1251; DB 1; Length 976;  
Best Local Similarity 32.7%; Pred. No. 1.4e-53;  
Matches 324; Conservative 166; Mismatches 352; Indels 148; Gaps 34;  
QY 47 VVKSSYPWSESPDLGALRQSSGTYEAAAEVDVYSASITLQVLVDAPNGISCLAW 106  
DB 20 VOTGSSQSPSPG-EFSPSTHFGKSDLI-----VRVDEIRLLCTDGFVK-WT 67  
QY 107 FKSSSLNCPHFEDLQNRGVVSMVILKMTETQAGEYLLFIQSEATN---YTILFTVSIRNT 163  
DB 68 FE-----ILDETENKQEWIT-EKAEATNTGKVTCTNKHGLSNS 106  
QY 164 L-----LYTLRRPYPRKMNODALVCIESVPEPIVWVLCDSQGESKEE----- 209  
DB 107 IYVVRDPAPKLELVDRSLYGK-EDNTLVRCPLTDPE-VTNYSKGCOCGKPLPKDLRFIP 164  
QY 210 ---SPAVKKEEKVLEHFGTDIRCCARNELGRECFRLFIDL---NOTPQTLFQ--L 260  
DB 165 DPKAGIMIKSVKRAVHRLC---LHCSVDQEGKSVLSEKFLKVRPAFAKVPVVSVKASY 221  
QY 261 FLKVGPELWIRCKAVHVNHGFLTWELN---KALEEGNYFEMSTYSTNRTMIRILFAFV 317  
DB 222 LLREGSEFTVTCTIKDVSSVYSTWKNRENSOTKLOEKYNSWHHGDENYERQAT---LTI 277  
QY 318 SSVARNDTGYTCSSSKHPSQSALVT---IVGKGFNA-----TNSSEDEYDOYE 365  
DB 278 'SSAVNDSGVFMCIYANTFGSANVTTLLEVVDKGFNIIFPMINTTVFVNDGENVDL---- 333  
QY 366 EFCFSVRFKAYPQ-IRCTWTFSSKSF--PCEOKGLONGYSISKFNHKK-----OPGE 415  
DB 334 ---IVEYEAFFKPEHQOYIMNRTFDKWEDYPKSENEINIRYSELHLHLRLKLGTEGGT 389  
QY 416 YIFHAENDAAQFTKMTFLNIRKRPQVLA--EASASQASCSFGDGYPLPSPWTKKCSOKSPN 473  
DB 390 YFVLNSDNDVAIAINNVYNTKPEILTYDLRVNGLQCVAAAGPEPTIDWYFCPGTEQR 449  
QY 474 CTEITE-GVMNRKANRKYVFGQWSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLNSP 532  
DB 450 CSASVLPVDVQTLNNSGPPGKLVQSSIDSSAFKNGTVECKAYNDVGKT--SAYFN-- 505  
QY 533 GPFPFIODN-----ISFVATIGVCLLFTLVLTLLCHIKYKQFRYESOLOM 578  
DB 506 --FAFKNKQEQTHPHTLFTPLLIGFVIVAGMMCIIMWILT---YKLOKPMVEVQWKV 559  
QY 579 V-OVTGSSDNEYFYVDFREYEDLKWEFFRENLEFGLGSGAFKGYMNATAYGISKTGV 637  
DB 560 VEEING---NNYVYIDPTQLPYDHKWEFFPNRLSFGKTLGAGAGKVVVEATAYGLIKSDA 616  
QY 638 SIQAVYMLKEKADSSEREALMSELKMMTOLGSHENIVNLLGACTLSGPIYLIPEYCCYG 697  
DB 617 AMTVAVKMLKPSAHLTEREALMSELKVLVSLGNHMINVLLGACTIGGPTLVITEYCCYG 676  
QY 698 DLLNLYLSRKEKFRHWTE-----IFKEHNFSPYPTFQSHPNSSM---PGSREVOIHPD 748  
DB 677 DLLNLRKRDSTFCSQEDAEALYKLLHLSKSSCSDSTNEMDMKPGSVTV-VPTK 735  
QY 749 SDOIISGLHNSHSEDEIEYENKRLREEEDLNVLTEDELLCFAYQVAKGMFELEKSCV 808  
DB 736 ADKRRSVRIGSY-----IERDVTPTAIMEDEL-ALDELLLSFSYQVAKGMFAFLASKNCI 789  
QY 809 HRDLAARNVLVTHGVKVICDFGLARDIMSDSNVYVRGNARLPVKWMAPESLFEGITIK 868  
DB 790 HRDLAARNILLTHGRITKICDFGLARDIKNDSNVYVKGARLPVKWMAPESEIFNCVTFE 849  
QY 869 SDVMSYGIILWEIFSLGVNYPGIPVDANFYKLLQNGFKMDQPFYATEEYIIMQSWAF 928  
DB 850 SDVMSYGIILWEIFSLGVNYPGIPVDANFYKLLQNGFKMDQPFYATEEYIIMQSWAF 928  
QY 929 DSRKRPSPNLTSLGCOLADAEAEAMQNV 958  
DB 910 DPLKRPFTQVQLVIEKQISESTNHHIYSL 939

**RESULT 8**

151703  
c-kit-related kinase 1 (Xkrkl) - African clawed frog  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 18-Jun-1999  
C;Accession: I51703  
R;Baker, C.V.; Sharpe, C.R.; Torpey, N.P.; Heasman, J.; Wylie, C.C.  
Mech. Dev. 50, 217-228, 1995  
A;Title: A Xenopus c-kit-related receptor tyrosine kinase expressed in migrating stem ce  
A;Reference number: I51703; MUID:95344996; PMID:7619732  
A;Accession: I51703  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-954 <BAK>  
A;Cross-references: EMBL:Z48770; NID:g763033; PIDN:CAA8688.1; PID:g763034  
C;Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
C;Keywords: Atg  
F;575-915/Domain: protein kinase homology <KIN>  
F;583-591/Region: protein kinase ATP-binding motif

Query Match 23.2%; Score 1226; DB 2; Length 954;  
Best Local Similarity 33.1%; Pred. No. 2,2e+52;  
Matches 305; Conservative 162; Mismatches 322; Indels 132; Gaps 30;

QY 107 FKHSSLCQPHEDLQRNVVS-----MVILKMPTQAGEYLFLFIQSSEATNYTLFTVSI 160  
I : | : ||| : | : | : | : | : | : | : | : | : | : | : | : | :  
DB 52 FOKSGLMKKPR-DLKSRPLNNSETDQFFVIKADLRHIGRYICTNTQTQENTSV--SLFV 108  
  
QY 161 RNTLLYTLPFRPKMENQDALVCISSEPPIPVWVLCDSQGESCKE-----ESPA 212  
:: : | : | : | : | : | : | : | : | : | : | : | : | : | :  
DB 109 KDAPRPFLDIPTIDVEGADVGMCFPTDPDMDTIAIEKCD-GSPLENFTTTOIEAGI 166  
  
QY 213 VVKKEKVLHELFGTDIRCCARNELGR-ECTRFLTIDLNPOTITLPOLFL-----KV 264  
: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
DB 167 TIKTVQLAFDSY-----VCSGNKKSTFSIHVPVK-KVPTVFUSKSQSLVKT 220  
  
QY 265 GEPLWRCKAVHNHGFLTW-ELENKALBEENGFYMSTYSTNRMTIMRLFAFVSSVARN 323  
||| : | : | : | : | : | : | : | : | : | : | : | : | : | :  
DB 221 GEPEVTCAVLDFESTVKAOWLVDVKGVTQAQNFRSSNFVSYNLTLS-----DGPPYS 274  
  
QY 324 DTGYTCSSSKPQS---SALTIVTKGFNAT-NSSDEYIDQYEERFCFVSVRFYAYPO- 378  
:: : | : | : | : | : | : | : | : | : | : | : | : | : | :  
DB 275 ESRTFCOAENAIGQVNATFTLDVIDGVYNLVLENTTISVNAGDNLVLYVIDAYPHP 334  
  
QY 379 IRCWTFSRKSPCEQKGLDNGYSISK-----PCNHK-----OPGEIFHAENDDA 425  
|| : | : | : | : | : | : | : | : | : | : | : | : | : | :  
DB 335 DDGWVTFYNETL----LNTSDHVATKDGENNRVSELHLRLKGTGKVTFYTYTNSDD 390  
  
QY 426 QFTKNFTLINRRKPQVLAPASOAS--CFSDGYPLPSWTWKCKSDKSNCTEEITEGYW 483  
: | : | : | : | : | : | : | : | : | : | : | : | : | :  
DB 391 DASVSFNIQVKTRPEIIAERTSEGTLQCVA TGFPVPAIQWFPGSGEORCTDYPLSPV 450  
  
QY 484 NRK--ANRVFCQWSSSLTNNSEAIKGLFKCCAYNSLGTSCTILLNSGPFFF-ID 540  
| : | : | : | : | : | : | : | : | : | : | : | : | : | :  
DB 451 NEKFIOENSLSGRVVVESTIDVNDLNKNGVQCVASNEVESAYSV-----FSAIKE 502  
  
QY 541 N-----ISFYATTGCCLLFWTLTLLICHKYKKQPRYESOLOMW-QVTGSSDNE 588  
| : | : | : | : | : | : | : | : | : | : | : | : | : | :  
DB 503 KLRTHTLFETPLLIGFTAAGLMCIAVAL----MYKLQRPKEYLOWKVVFEING---NN 555  
  
QY 589 YFYVDREYEDLKWEPRENLEFFGKVLGSAGFKVMNATAYGISTGVSIOQVAKMLKE 648  
| : | : | : | : | : | : | : | : | : | : | : | : | : | :  
DB 556 VYVIDPTQLPYDNKNWEFFRDRLCFGILGAGAFKVVEATAYGLLKEDSRULTVAVKMLKP 615  
  
QY 649 KADSSERALMSKLMMTQLGSHENVNLLGACTLSGPYLIIFYECYGDLLNLYRSKR 708  
| : | : | : | : | : | : | : | : | : | : | : | : | : | :  
DB 616 SAHSTEREALMSKLVSYLGHHKNLVNLLGAC TVGGPTLVITYCYCGDLLNLYRRKRD 675  
  
QY 709 KFHRTWTEIFEKHNHSFYTFEQSHPNs---SMGPSREVQHHPDSQDI----- 752  
DB 676 SF-----ICPKFDENSAALKYKLNLLNTRMDMGCEGMSEYIDMKPAVSYYVPT 721  
  
QY 753 --SGLHGNFSHSEDIEYENQKREEEDLNVLTFEDLLCFAYOVAKGMEFLFKSCVHR 810

[illegible]

**RESULT 8**

151703  
c-kit-related kinase 1 (Xkrk1) - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C>Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 18-Jun-1999  
C:Accession: I51703  
R:Baker, C.V.; Sharpe, C.R.; Torpey, N.P.; Heasman, J.; Wylie, C.C.  
Mech. Dev. 50, 217-228, 1995  
A>Title: A Xenopus c-kit-related receptor tyrosine kinase expressed in migrating stem ce  
A:Reference number: I51703; MUID:95344996; PMID:7619732  
A:Accession: I51703  
A>Status: preliminary; translated from GB/EMBL/DDBB  
A:Molecule type: mRNA  
A:Residues: 1-954 <BAK>  
A:Cross-references: EMBL:J48770; NID:g730333; PIDN:CAA8688.1; PID:g763034  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology:  
C:Keywords: ATP  
F:575-915/Domains: protein kinase homology <KIN>  
F:593-591/Region: protein kinase ATP-binding motif

Query Match 23.2%; Score 1226; DB 2; Length 954;  
Best Local Similarity 33.1%; Pred. No. 2.2e+52;  
Matches 305; Conservative 162; Mismatches 322; Indels 132; Gaps 30;

QY 107 FKHSLSNQPHFDLQRGVVS-----WVILKTQTAGEYLFLFIQSATNTIIFTYSI 160  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 52 FOKSLMKPKPR-DLKSRPLNNSTDOFFVIKADLRHIGRYICTNETQENTSV--SLFV 108  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 161 RNTLLYTLLRRPVPFRKMENODALVCISESVEPIVENVLVCDSQGESCKE-----ESPA 212  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 109 KDPARPFIDIPDVTGEGADVGMCFPTDPMDAIEKD--GPSLPENFTTDTIEAGI 166  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 213 VVKBEKVHLHFGLTDIRCCARNELGR-ECTRFLTIDLNPQTTLPOLFL-----KV 264  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 167 TIKTVQLAFDSQCY----VCSGNKSCTGVKKSSFSIHVKPVKP-KVPTVFLSKSQLVKT 220  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 265 GEPLMIRKAVHHNGFGLTW-ELENKALEEGNVFMSTYSTNRMTIRILFAFVSSVARN 323  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 221 GEPEFVTCAVLDFVSTVKAQWLVDKEGTVKQANPRSSNVFSYNLTLS-----DGVPYS 274  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 324 DTGYTCSSSKHPQ---SALTIVKGGFNAT--NSSEDYEDIQYEERFCFVRFKAYPO- 378  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 275 ESRTFOQAENAIGNATFTLDVIDGVYNLTVLENTISVNAGDNLVKLVIIDAYPHP 334  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 379 IRCWTFSRKSPCEOKGLDNGYSISK-----PCNHKH-----QPGEYIFHAENDDA 425  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 335 DDGVWTFNYETL----LNTSDHYVATKDEGNRRVSELHLRLKLGTEKGVTYFITNSDD 390  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 426 QTKMFTLNIRKPKOVLAEASQAS--CFSDGYPLPSWTWKKCDKSPNCTETEITGVW 483  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 391 DASVSFNQVKTTRPELIAERTSEGLTQCVAATFPVPAIOWFYFCPGSEQRCRDTPPLSPV 450  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 484 NRK--ANKRVFGOWSSSTLNMSSEAIGFLVKCCAYNSLGTSCTETILLNSGPFPF-ID 540  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 451 NEKFQNSSLSGRIVVESTIDVNDLKNKTVQCVASNEVESAYSV-----FSFAIKE 502  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 541 N-----ISFYATIGVCLLFVVTLFLLCCHKYKQPRYESQLMW-QVTGSDDNE 588  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 503 KLRTHTLFTPLLIGFTAAGLMCIAVAL---MYKLOKPKYELQWKVVEING-----NN 555  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 589 YFYDVREYEDLKWEPRENLEFFGKVGSGAFGKVMNATAYGISTGVSIOQVAYMKLKE 648  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 556 YVYIDPTQLPYONKWEPPDRDLFCFGLGAGAFGVVYEAATAYGLLKEDSRILTVAVKMLKP 615  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 649 KADSSERALMSCLKMMTOIGSHENIVNLGCATLSGPYLIIFYCCYGDLLNIURSKRE 708  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 616 SAHSTEREALMSCLKVSYLGHHKHNVLNLAGCATVGTTLVITEYCXYGDLLNLYRRKR 675  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 709 KFHRTWTEIFEKHNFSSFYPFEQSHPN-----SMPGSREVQHHPDSQDI----- 752  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 676 SF-----ICPKFDNSAAALKYKLNLLNTRDMCGMGMEYIDMKPAVSYVVPT 721  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
QY 753 --SGLHGNSFHSEBIEYENQNKRLEEEDLNVLTFEDLLCFAYOVAKGWEFLEFKSCVHR 810  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY	289	NKALEGN----	YFEMSTYSTNRTMIRILFAFVSSVARNDGTGYTCCSSK----	HPQSQAL	344
DB		: :	: :	: :	
DB	236	--SLRHGDTKLITISQSDFDHNEYQ--	KVLTLNLDHVSFQDAGNYSCTATNAGNHSASMV	292	
QY	342	VTIVGKGFINATN--	SSEDYEIDQYEFECFSVRKAYPQIRCT--	TWTSRKSFCPEQKGLD--	398
DB		: :	: :	: :	
DB	293	FRVVEAYLNLTSQESLLQEVTVGKEVLDQVKVEAYPGLESFNWY--	LGPFSDYQDKLDF	351	
QY	399	-----NGYSISKFCN----	HKHGPKEYIFHAENDDAQFTKMTLNIRKKPQVLAEASASQA	450	
DB	352	VTIKDYRYTSTLSPLRKLKSEAGRYSFLARNAGGQNALTFELTLYR	PEVRYVTMLING	411	
QY	451	S-----CFSDGYPLSPWTWKKCDKSPNCTEE--	ITEGVNMRKANRKYFGQVQWSSSTLNM	503	
DB	412	SDTLCEASGYPQPSVTWVQCRSHTRDCRDE	SAGLVLEDHSDSEVLSSQVPFPEVIVHSLLA	471	
QY	504	SEAIKGLVKCCAYNSLGTSCETILLNSPGPPFPFIQDNISFYATIGVC--	LLFIVVLTL	560	
DB	472	GTLHENRTYECRAFNSVGNSSQFTWPISGAHTOLPDELFTPVLT	CMISIMALLULLLL	531	
QY	561	LICHYKKQFRYESOLOMVOVTVSSDNEYFYVDFREYEDLKWEFFPRENL	EFGKVLGSGA	620	
DB	532	LLLTKYKQKPKYQVRKILIE--	SYEGNSYTFIDPTQLPYNEKWEFFPRNNLFGKTLGAGA	589	
QY	621	FGKYMNATAYGISKTGYSIQVAVKMLKEKADSSREALMSSELKMTOLGSHENIVNLLGA	680		
DB	590	FGKVVEATAFGLGREDAVLKVAVKMLKSTAHADKEALMSSELKIMSHLGOHENIVNLLGA	649		
QY	681	CTLSGPIYLFEYCCYGDLLNLYLSKRE-----	KPHRTWTEIFKE	720	
DB	650	CTHGGPVLVITEYCCYGDLLNFLRQAEAMLGPSLVGQDPEAGAGYKNHLEKKYVVRD	709		
QY	721	HNFSFYPTFQSHNPMSPGSRREVQIHDPDSQISLHGNSEFHSDEIEYENOKRLEEEDL	780		
DB	710	SDFS-----	SGQVTVYEMRPVSTSSNSDSFSEEDLCKGDRPLE-----	749	
QY	781	NVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLARNVLVTHGKVKVICTDGLARDIMSDS	840		
DB	750	-----LRDLLHFSSQAQGMALASKNGCIHRDVAARNVLLTSGRVAKIGDGLARDIMDS	805		
QY	841	NYVVRGNARLPVKWMAPESLFEGYTTKSDVWSYGILLWEIFSLGVNYPYGPIDVANDFYK	900		
DB	806	NYIVKGNARLPVKWMAPEISFDCVTVQSDVWSYGILLWEIFSLGOLNYPYGPILLVNSFYK	865		
QY	901	LIONGKMDQPFYATEEYIIMQSCWAPDSRKRPSFPNLTSLFCQL--	ADAEAEAMYQNV	958	
DB	866	LKVDGYQMAQAPAFAPAKNIYSIMQACWALEPRTPTFQICSLLKQQAQEDRRVPNYTNL	924		
RESULT 10					
TVWVMD					
protein-tyrosine kinase (EC 2.7.1.112) fms precursor - feline sarcoma virus					
C:Species: feline sarcoma virus					
A:Note: host Felis sp. (cat)					
C:Date: 27-Nov-1985.#sequence_revision 31-Dec-1991 #text_change 13-Jun-1991					
C:Accession: A00654					
R:Hampe, A.; Gobet, M.; Sherr, C.J.; Galibert, F.					
Proc. Natl. Acad. Sci. U.S.A. 81, 85-89, 1984					
A:Title: Nucleotide sequence of the feline retroviral oncogene v-fms showing					
A:Reference number: A00654; MUID:84119469; PMID:6582485					
A:Accession: A00654					
A:Molecule type: DNA					
A:Residues: 1-941 <HAM>					
C:Comment: This protein is synthesized as a gag-fms polyprotein.					
C:Genetics:					
A:Gene: fms					
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin					
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transfer					
rotein kinase					
F:1-23/Domain: signal sequence #status predicted <SIG>					
F:24-941/Product: protein-tyrosine kinase fms #status predicted <MAT>					
F:24-509/Domain: extracellular #status predicted <EXT>					
F:35-96/Domain: immunoglobulin homology <IMM>					



Db 560 SYTFIDTQPLPYNEKWEFFPRNNLQFGTLGAGAEKGVVETAFGLGKEDAVLKVAVKMLK 619  
Qy 648 EKADSSREALMSKMMTOLGSHENINVLGACTLSGPIYLIFECYCYGDLNLYLSRKR 707  
Db 620 STAHADKEALMSKIMSHLQHENINVLGACTHGGPVLVITCYCYGDLNLYLSRKA 679  
Qy 708 E-----KFWRTWTFEKEHNFSEYPTFFQSHPNSSMPGSRVQIHP 747  
Db 680 EAMLGPSLSPQDPEGVDYKNIHLKRYVRDGSF-----SQGVDTYV 724  
Qy 748 DSDOISGLHNSHSEDEIEYENOKRLEEBDLNVLTFEDLLCFAYOVAKGMELEFKSC 807  
Db 725 EMRPVSTSSDSF--SEODLKEDGRPLE-----LRDLLHSSQVAQMAFLASKNC 774  
Qy 808 VHRDLAARNVLTHGKVVYKICDGLARDIMSDSNVVRGNARLPVKMMAPESLFEPIY 867  
Db 775 IHRDVAARNVLLTNGHAKIGDFGLARDIMSDSNVVRGNARLPVKMMAPESIFDCVTV 834  
Qy 868 KSDVWSYGILLWEIFSLGVNPGIPVDANFYKLIQNGFKMDQPFYATEEYIIMOSCA 927  
Db 835 QSDVWSYGILLWEIFSLGVNPGIPVDANFYKLIQNGFKMDQPFYATEEYIIMOSCA 894  
Qy 928 FDSKRKPSFNLTSFLGCOL-ADAEAMQYQV 958  
Db 895 LEPHRPTFOQICSFLOBAQEDRRERYTNL 926  
RESULT 12  
JN0677  
protein-tyrosine kinase (EC 2.7.1.112), receptor type kit precursor - chicken  
N:Alternate names: tyrosine kinase receptor kit  
C:Species: Gallus gallus (chicken)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
C:Accession: JN0677  
R:Sasaki, E.; Okamura, H.; Chikamune, T.; Kanai, Y.; Watanabe, M.; Naito, M.; Sakurai, M.  
Gene 128, 257-261, 1993  
A:Title: Cloning and expression of the chicken c-kit proto-oncogene.  
A:Reference number: JN0677; MUID:9329295; PMID:7685729  
A:Accession: JN0677  
A:Molecule type: mRNA  
A:Residues: 1-960 <SAS>  
A:Cross-references: DDBJ:D12225; NID:g303532; PIDN:BAA02506.1; PID:g303533  
A:Experimental source: brain  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;  
C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;  
rosetin kinase  
F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-960/Product: tyrosine kinase receptor #status predicted <MAT>  
F:314-380/Domain: immunoglobulin homology <IMM>  
F:573-916/Domain: protein kinase homology <KIN>  
F:581-589/Region: protein kinase ATP-binding motif  
F:76,135,149,269,286,306,318,338,356,453,469/Binding site: carboxydrate (Asn) (covalent)  
F:76,135,149,269,286,306,318,338,356,453,469/Binding site: carboxydrate (Asn) (covalent)  
Query Match 23.0%; Score 1215; DB 1; Length 960;  
Best Local Similarity 32.3%; Pred. No. 7.5e-52;  
Matches 314; Conservative 165; Mismatches 370; Indels 124; Gaps 34;  
Qy 66 ALRPOSSGTVEAAVEVDVSNASITVLQVLDAPGNITSLWPKHSS-----LNCOPHFD 119  
Db 18 SLIPAGGVPHEESSVLKGEELKCNKEGP-----VTWTFQNDSPSAKTRISNEKEWH 73  
Qy 120 LQNRGV--VSMVLKMTQTQAGEYLLFIOSEATNTYTLFTVIRNTLLYLRPRFRKME 177  
Db 74 TKNATIRIGRECKSKGSIVNSFYVFKDP---NVLFLV---DSLIV-----GKED 119  
Qy 178 NODALVCTSESVEP-IVEWVLCDQSGESCKEESPAVVKKE-----KVLHFLFGTDIRCC 232  
Db 120 SDILLVC---PLTDPDLNFTLRKCDGKPLPKMTFIPNPKGIIKKNVQSRFKGICYQL 176  
Qy 233 AR-NELGRECTFLFDLNLQTP-QTTLPOL-----FLKVGEPWLIRCKAVVNHGFL 283  
Db 177 AKHNGVEKISEHIF---LNVVRVHPALPITLSKSYELLKEGEEFEVTCITDVS SVKA 233

Qy 284 TWELNKALEEGNYFEMXSTYSTNRTMIRILFAFYSSVARNDTGYVTCSSKHP--SQSAL 341  
Db 234 SWISYKSAIVTSKRNLDGYERK-----LTLNRSVGUNDGSEFTC-QAENPFCKTAT 288  
Qy 342 VTI--VGKGFIN--ATNSSDEYEDQYEEFCFSYRFRKAYQIR-CTWTFESRKSPPCEOK 395  
Db 289 VTLKALAGFVRLFATMTTIDINAGNGN--LTVEYEAYPKPKKEEVMMYMETL---QN 343  
Qy 396 GLDNGYSTSKFCNKH-----OPGEYIHAENDDAQTKMFTLNRKPOVLA 443  
Db 344 SSDHYVKFKTVCNNSYTSSELHLRLKLGTEGGIYTFVNSDASSSVTFNVYVTKPEILT 403  
Qy 444 EASASQ--ASCFSQGYPLPSTWTKKCDKSPNCTEITEGVNMRKANRKFVQWVSSSTL 501  
Db 404 LDMIGNDILQCVATGFPAPTYIWFYFCPQTEQRCLOSPITSPMDVKVS-----YNSVYP 457  
Qy 502 NMSEAIGFLVKCCAYNSLGTSCETILLNSPGPPF---IODNISFYA-----TICV 550  
Db 458 SFERILVESTVNASMFKSTGTICCEASSNGDKSSVFENFAKEQIRHTLFTPLLIAFGV 517  
Qy 551 CLLFTVWLTLICHYKKQFYESQLOMV-QVTGSSDNEYFYVDFREYEDLKWEPPREN 609  
Db 518 AAGLCIIIVMLIVLYLQPKYEVQKVVVEING---NNVYVIDPTQLPYDHKWEPPNR 574  
Qy 610 LEFGKVLGSGAGKVMNATAYGISKTGYSVIOAVKMLKEKADSSREALMSLKMMTOLG 669  
Db 575 LSFGLTLAGAGFGKVVETATAYGLFKSDAAMTVAVKMLKPSAHLTEREALMSLKLVLSLG 634  
Qy 670 SHENVNLLGACTLSGPIYLIFECYCYGDLNLYLSRKRKF---HRTWTFEIKHEHNSF 725  
Db 635 NHINIVNLLGACTIGPTLVITEYCYCYGDLNLYLSRKRDSFCPKHEEHAERAVYENL-- 692  
Qy 726 YPTFOSHNS-----SMGPSREVOIHPDSOISGLHNSHSEDEIEYENOKRLEEE 778  
Db 693 --LHQAEPTADAVNEYMDKPGVAVPPKADKRPVKSGSYTDQD---VTLSMLEDE 746  
Qy 779 DLNVLTFEDLLCFAYOVAKGMELEFKSCVHRDLAARNVLTHGKVVYKICDGLARDIMS 838  
Db 747 --LALDVEDLLSFYQVAKGMSFLASKNCHRDLAARNILLTHGRITKICDGLARDIRN 804  
Qy 839 DSNVYVVRGNARLPVKMMAPESLFEPIYTKSDVWSYGILLWEIFSLGVNPGIPYVDANF 898  
Db 805 DSNVYVVRGNARLPVKMMAPESIFNCVTFESDWSYGILLWEIFSLGSSPYGMPVDSKF 864  
Qy 899 YKLQNGFKMDQPFYATEEYIIMOSCAFDKRPSPFNLTSLFCQLADAEAMQYQV 958  
Db 865 YKMIKEGYRMESPECSPPPEMYDIMKSCWDADPLQRTTFKQIVQLIEQQLSDNAPRYAN- 923  
Qy 959 DGRVSECPHYTON 971  
Db 924 ---FSTPPSTOQGN 933

RESULT 13  
T30816  
macrophage colony-stimulating factor receptor - Japanese pufferfish  
C:Species: Fugu rubripes (Japanese pufferfish)  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Jun-2000  
C:Accession: T30816  
R:How, G.F.; Venkatesh, B.; Brenner, S.  
Genome Res. 6, 1183-1191, 1996  
A:Title: Conserved linkage between the pufferfish (Fugu rubripes) and human genes for  
A:Reference number: 220882; MUID:97129405; PMID:8973913  
A:Accession: T30816  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-975 <HOW>  
A:Cross-references: EMBL:U63926; NID:g1752706; PID:g1752708; PIDN:AAC60063.1  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo  
Query Match 23.0%; Score 1213; DB 2; Length 975;  
Best Local Similarity 31.7%; Pred. No. 9.6e-52;

Matches 325; Conservative 177; Mismatches 333; Indels 190; Gaps 37;			
Qy	33	IKCVLINHKNNDSSVKGKSSYP-----MYSESPEDLGC--ALRP--QSSCT--VYEAARV	81
Db	47	LKC-----GGDPYVWQTRLPKHRYMSRSPGNLRTIRVAPTAFTYTYKCFYSARQA	100
Qy	82	EVDVSATITLQVLVDAPNISCLWVKHSSLNCPHFQDLNQRGVVSMVLKMTQTQAGY	141
Db	101	HRHLTSSV--HVYVDPNRV--FW-----TSSTSLRVVRKEGDDY	136
Qy	142	L--LFIOSEATNYTILTVSIRNFTLLYLRLRPYPRKMNQDALVCISEVPPIVWVL	198
Db	137	LLPCLLTDPEDATDGL-----RMDN-----GTVVP-PEMNYTV	168
Qy	199	CDSOGESCKEESPAVVKKEKVLHFGTDIRCCARNELGRECTRLFTIDLNQTPQTLIP	258
Db	169	YRHGILIRLSOPS-----FNADYVCTAKVGVEKTSFTFSINVLQKLRFP-P	215
Qy	259	QLFLK-----VGEPLMIRKAVHNVHGFGLTWELNKA-----LEEGNYFEMSTYSTR	308
Db	216	YVFLMEDEYVRIVGEELQIRCMTHNPENYNVWNTYTKSRVTIEE---RVRSSGENRL	271
Qy	309	MIRILFAFVSSVARNDTGYTCSSSKHP---SQSALVTIVGKFI-----NATNSSE	357
Db	272	DIQGLIT-ISAVDLADTGNISCICTNEAGVNSNTYLLVVEKPYIRLWLPQILPKLASOGL	330
Qy	358	DYEIDQVEEFCFSYRFRKAYPOIR-CTWTFSRKSPCEOKGLDNGYSISKFNCHKHQPGEY	416
Db	331	SVENEGEDLELGVNVEAYPOITDRHWHTPTSPSTSMGEIYHARLQKRWNAQEQ-QGY	389
Qy	417	IFHAENDDAQFTMTLNIIRKP-QVLAESAASQACSFSDGYPLPSWTWKCKSDKSPNCT	475
Db	390	TFYAKNSLANGSISFHVVMKYKPIAVRWENITLTCTSFYGPAPQIIWQCSGIRPTCN	449
Qy	476	EEITEGVNWR-----KANRKFVGQWVSSSTLMSAIGFLVKCCAYNSLGTSCETI	527
Db	450	GNNP-GLPKQNHQALTVQVOREEYGAVESEVTFVGLSNRMTVECFVNLGVSSDTF	508
Qy	528	LNSPGPPFPFIODNISFVATI---GVCLLFIIVLTLLICHYKKQRYESOLOMVOVTG	583
Db	509	TVE-----VSDKL-TSTSLIGAAGVLAIFELLVFLV-KYKQKPFERWKIIE--A	557
Qy	584	SSONEYPIVDFREYEDLKWEFPRENLEFGKVLGSGAFKVMNATAYGISK-TGVSIQVA	642
Db	558	REGNNTFIDPTQLPYNEKEWEPFRDKLKLGLVLAGAGAFKVVVEATAFGLGEDKONTLRA	617
Qy	643	VKMLKEKADSSEREAALSELKMTQOLGSHENIVNLGACTLSGPIYLIFFCYCGDLNY	702
Db	618	VKMLKANAHSDEREAALSELKILSHLGHQHNIVNLGACTYGGPVLYITEYCSLGLDNF	677
Qy	703	LRSKREKFRHTWTEIFREHNFSEYPTQSHFN-----SSMPGSRVQV	744
Db	678	LQKAETFNVLVMI-----PEIMENNDYKNCNQKWIIRSDSGISSTSSSTYLE	728
Qy	745	IHPDSQOISGLHGNFSFSEDEIYENOKRLEEBEDLNVLTEDLCCFAYQVAKGMFELEF	804
Db	729	MRPS-----QOSHIEASGRKSLCEDNGDNPDLDDLLRFSLOVAQGLDFLAS	775
Qy	805	KSCVHRLAARNVLVTHGKVKVVICDFGLARDIMSDSNVYVVRGNARLPVKWMAPSLFGI	864
Db	776	RNCIHRDVAARNVLLTKRVAKICDFGLARDIMSDSNVYVVRGNARLPVKWMAPIESIFCV	835
Qy	865	YTKSDVMSYGILLWEIFSLGVNYPGIPVDANYKLIQNGCFKMDQDPYATEEYIIMOS	924
Db	836	YTVOSDVMYSYGILLWEIFSLGVNYPGIPVDANYKLIQNGCFKMDQDPYATEEYIIMOS	895
Qy	925	CWAFDSKRKSPFNLTSLFLGCOLA---DAEAMYNV-----DGRVSECPTHYQNRPF	975
Db	896	CWNLEPTEPTFMSISQMINRLCGQDEQEKLIYRNVOPEQVAGEACDEPKRYD--PPC	953
Qy	976	SREMD 980	
Db	954	ERSCD 958	

RESULT 14  
S16385  
macrophage colony-stimulating factor 1 receptor precursor - rat  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) CSF-1R  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 02-Jun-2000  
C:Accession: I60321; S16385  
R:Borycki, A.G.; Guiller, M.; Leibovitch, M.P.; Leibovitch, S.A.  
Growth factors 6, 209-218, 1992  
A:Title: Molecular cloning of CSF-1 receptor from rat myoblasts. Sequence analysis an  
A:Reference number: I60321; MUID:93001225; PMID:1389227  
A:Accession: I60321  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-978 <RES>  
A:Cross-references: EMBL:X61479; NID:g57543; PIDN:CAA43706.1; PID:g57544  
A:Note: in Genbank entry RRCSF1, release 113.0, the source is designated as Rattus ra  
A:Note: submitted to the EMBL Data Library, August 1991  
C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homolo  
C:Keywords: ATP; autophosphorylation; glycoprotein; growth factor receptor; kinase-re  
protein; tyrosine-specific protein kinase  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-978/Product: macrophage colony-stimulating factor 1 receptor #status predicted <  
F:20-515/Domain: extracellular #status predicted <EXT>  
F:35-86/Domain: immunoglobulin homology <IMM1>  
F:120-179/Domain: immunoglobulin homology <IMM2>  
F:1217-280/Domain: immunoglobulin homology <IMM3>  
F:316-381/Domain: immunoglobulin homology <IMM4>  
F:410-485/Domain: immunoglobulin homology <IMM5>  
F:516-535/Domain: transmembrane #status predicted <TM>  
F:536-978/Domain: intracellular #status predicted <INT>  
F:578-915/Domain: protein kinase homology <KIN>  
F:586-594/Region: protein kinase ATP-binding motif  
F:42-84,127-177,224-278,417-483/Disulfide bonds: #status predicted  
F:45,73,302,335,389,410,449,478,491/Binding site: carbohydrate (Asn) (covalent) #stat  
F:614,631,776/Active site: Lys, Glu, Asp #status predicted  
F:781,794/Binding site: magnesium (Asn, Asp) #status predicted

Query Match 22.6%; Score 1194; DB 2; Length 978; Best Local Similarity 33.1%; Pred. No. 8e-51; Matches 321; Conservative 154; Mismatches 332; Indels 164; Gaps 32;			
Qy	83	VDVSASITLOVL---VDAPGNISCLWVKHSSLNCPHFQDLNQRGV-----SWILKM	133
Db	32	VEPGETVTLRCVNSGVEMDGPISPYWTLDPES---PGSTLTTRNATFKNTGYRCTEL	87
Qy	134	TETOAGEYL--LFIOSEATNYTIL---FTVSIKNTLLYTLRLPYPRKMNQDALVCISES	188
Db	88	EDPMAGSTTIHLVYKDPASHWMLLAQEVTV-----VEQEA-----	124
Qy	189	VPEPIVWLCDSDGESCSESPAVVKK-----EEKVLHFLFGTDIRCCAR	234
Db	125	LPCLITDPAKDVSV-SLMREGGQVLRKTVYFFSAWRGFIIRKAKVL-----DSNTVCKT	179
Qy	235	NELGRECTRL-FTIDLN---QTPOTTL-PQLFLKY-GEPLMIRKAVHNVHGFGLTWEL	287
Db	180	MVNGRESTSTGTLKVRNVHVPPEPQIKLEPSKLVIRGEAAQIVCSATNAEYGVNILKR	239
Qy	288	ENKALE-----EGNYFEMSTYSTNRTMIRILFAFVSSVARNDTGYTC-SSSKHPSQS	339
Db	240	GDTKLEIPLNSDFQDNYK-----KVRALSLNAVDFQDAGIYSCVASNDVGT	288
Qy	340	ALVT--IVKGFINATN-SSDEYEDQVEEFCFSYRFRKAYPOIR-CTWTFSRKSPCEOK	395
Db	289	ATMNFQVVEAYLNLTSQSLLOEVSQSLTLTVHADAYPSIQHYNNWYLGPFPE-DQR	347
Qy	396	GLD-----NGYSISKFCNH--KHQPGEYIFHAENDDAQFTKMTLNIIRKRPQVLAES	446
Db	348	KLEFITORAIRYRTFKFLNRVKASEAGYFLMAQNKAGWNNLTETLTRYPEVSVTWM	407
Qy	447	ASQAS-----CFSGDGYPLPSWTWKCKSDKSPNCTEITEGVN-----RKANRKFVQGVSS	498

Db 408 PVNSDVLFCDSVGPQPSVTWMECRGHTDRCDQAALQVNDTHPEVLSQKPFQKVIQ 467  
Qy 499 STLNMSEAIGFLVKCCAYNSLGSCEPILNSPGPPFIQDNISFYATIGVC---LLFI 555  
Db 468 SOLPIGLTKHNTYFCKTHNSVGNSSQYFRAVSLGQSKQLPDESFTFPVVACMSVMSLL 527  
Qy 556 VYLLILCHYKQFRIESQLOMVOVTGSSDNEYFYVDREYEDLKWEEFRENLEFCVK 615  
Db 528 VLLLLLLLYKQKPKYQVRKIIIE--RYEGNSYTFIDPTOLPYNEKWEFPRNNLOFQK 585  
Qy 616 LGSAGFQVMNATAYGISTGVSIOAVKMLKEKADSSEREALMSLKMMTOLGSHENIV 675  
Db 586 LCAGAFKVVATEAGLGKEDAVLKVAVKMLKSTAHADKEALMSLKMIMSHGOHENIV 645  
Qy 676 NLLGACTLSGPIYLLFYCCYGDLLNLYRSKRE-----KFHRTWT 715  
Db 646 NLLGACTHGGPVLVITEYCCYGDLLNFLRRKAEAMLGPSLSPGQDSEGDSSYKNHLEKK 705  
Qy 716 EIFKEHNFSPYTFQSHNSNMPGSRVQIHPDSDQISLHGNPSHSEDEIEYENQKRL 775  
Db 706 YVRDGSF-----SQGVDYVEMRPVSTSSDFFKQD-LDKEPSRPLE 749  
Qy 776 EEDLNVLTFEDLLCFAYQVAKGMEFEFKSCVHRDLAARNVLVTHGKVKICDFGLARD 835  
Db 750 -----LWDLHFSSQVAGMAFLASKNCIHRDVAARNVLLTSGHVAKIGDFGLARD 800  
Qy 836 IMDSNYYVRGNARLPVKWMAPEFEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVD 895  
Db 801 IMDSNYYVRGNARLPVKWMAPEFEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVD 860  
Qy 896 ANFYKILONGFKMDOPFATEIYIIMQSCWAFDSKRPSPNLTSLFCQLADAEAMY 955  
Db 861 NKFYKLVKDGVOQAOPVAPKNYISIMQSCWDLPTRRPTFOQICFL-----QEQARL 914  
Qy 956 QNVDCRVSECP 966  
Db 915 ERRDQDYANLP 925

## RESULT 15

## TWMSND

Macrophage colony-stimulating factor 1 receptor precursor - mouse

N:Contains: protein-tyrosine kinase. (EC 2.7.1.112) csflr/fms

C:Species: Mus musculus (house mouse)

C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 02-Jun-2000

C:Accession: S01880

R:Rothwell, V.M.; Rohrschneider, L.R.

Oncogene Res. 1, 311-324, 1987

A:Title: Murine c-fms cDNA: cloning, sequence analysis and retroviral expression.

A:Reference number: S01880; MUID: 88217329; PMID: 2966922

A:Accession: S01880

A:Molecule type: mRNA

A:Residues: 1-976 &lt;ROT&gt;

A:Cross-references: EMBL:X06368

C:Genetics:

A:Gene: fms

C:Superfamily: macrophage colony-stimulating factor 1 receptor; immunoglobulin homology;

C:Keywords: ATP; autophosphorylation; glycoprotein; kinase-related transforming protein;

fic protein kinase

F:1-19/Domain: signal sequence #status predicted &lt;SIG&gt;

F:20-976/Product: macrophage colony-stimulating factor 1 receptor #status predicted &lt;MAT

F:20-515/Domain: extracellular #status predicted &lt;EXT&gt;

F:35-86/Domain: immunoglobulin homology &lt;IMM1&gt;

F:120-179/Domain: immunoglobulin homology &lt;IMM2&gt;

F:217-280/Domain: immunoglobulin homology &lt;IMM3&gt;

F:316-381/Domain: immunoglobulin homology &lt;IMM4&gt;

F:410-485/Domain: immunoglobulin homology &lt;IMM5&gt;

F:516-535/Domain: transmembrane #status predicted &lt;TM&gt;

F:536-976/Domain: intracellular #status predicted &lt;INT&gt;

F:578-914/Domain: protein kinase homology &lt;KIN&gt;

F:586-594/Region: protein kinase ATP-binding motif

F:42-84,127-177,224-278,417-483/Disulfide bonds: #status predicted

F:45,73,302,335,389,410,449,478,491/Binding site: carbohydrate (Asn) (covalent) #stat  
E:614,631,776/Active site: Lys, Glu, Asp #status predicted  
F:781,794/Binding site: magnesium (Asn, Asp) #status predicted

## Query Match

22.4%; Score 1183.5; DB 1; Length 976;

Best Local Similarity 33.2%; Pred. NO. 2.6e-50;

Matches 318; Conservative 154; Mismatches 347; Indels 139; Gaps 30;

Qy 83 VDVASITLQVL-----VDAPGNISCLWVF---KHSNLNCPHFIDLQNRGVSMVLKMTF 135  
Db 32 VEPGETVTLRCVNSGVEMDGPISPIWLDLPSPGSTLTTSNATFNKTYRCTLEDDPM 91

Qy 136 TOAGEYLLFIOSEATNYTL---FTVISRNTLLYLRPRYFRKMEQNALVCISESVP 192  
Db 92 AGSTIHLVVRDPAHSNLLAQEVTV-----VEQGEAV-----LPCL 128

Qy 193 IVEWVLCDSQGESCKEESPAVYKKEKVLHFLFGTDIR-----CCARNELORECT 242  
Db 129 ITDPALKDSV-SLMREGGRQVLRKTVYFFSPWRGSIIRKAKVLDNSTYVCKTMVNGREST 187

Qy 243 RL-FTIDLN---QTPQTL-PQLFLKV-GEPLWIRCKAVHVNHGFLTWELNKALE-- 293  
Db 188 STGIWLKVNVRHPEPQIKLEPSKLVIRGEAAQIVCSATNAEVGFNVILKRGDTKLEIP 247

Qy 294 -----EGNYFEMSTYSTNTRMIRILFAFVSVARNDTGYYTC-SSSKHPSSQALVT--IV 345  
Db 248 LNSDFQDNYK-----KVRALSNAVDFODAGIYSCVASNDVGTTRATMNFQV 296

Qy 346 KGKFINATN-SEDEYIDQYBEFCFSVRKAYPQIR-CWTFSRKSFPCEQKGLD----- 398  
Db 297 ESAYLNTSEQSLLOEVSVDLSLITVHADAPSIQHYNWTYLVGPPEF-DQRKLEITOR 355

Qy 399 --NGYSISKFCNH--KHQGEYIFHAENDDAQFTKMTLNIRKPOVLAESAQAS--- 451  
Db 356 AIYRYTEKFLNLRVAKASEAGQYFLMAQNKAGNNLTFTLTLYPPEVSVTWMPVNGSDVL 415

Qy 452 -CFSDGYPLPWSWTWKCKSDKSPNCTEETIEGVN-----RKANKRVGQVSSSTLNSEA 506  
Db 416 FCDVSGYPOPSVTWMECRGHTDRCDQAALHLNWDTHPEVLSQKPFQKVIQSLPIGL 475

Qy 507 IKGFLVKCCAYNSLGTSCETILLNSPGPPFIQDNISFYATIGVC---LLFIYVLLIIC 563  
Db 476 KHNMTYFCKTHNSVGNSSQYFRAVSLGQSKQLPDESFTFPVVACMSVMSLLVLLLL 535

Qy 564 HKYKQFRIESQLOMVOVTGSSDNEYFYVDREYEDLKWEEFRENLEKVLGSAFGK 623  
Db 536 YKYKQPKYQVRKIIIE--RYEGNSYTFIDPTOLPYNEKWEFPRNNLOFQKTLGAGAFGK 593

Qy 624 VMNATAYGISTGVSIOAVKMLKEKADSSEREALMSLKMMTOLGSHENIVNLLGACTL 683  
Db 594 VVEATAFGLGKEDAVLKVAVKMLKSTAHADKEALMSLKMIMSHLOHENIVNLLGACTH 653

Qy 684 SGPIYLIFCYCCYGDLLNLYRSKREKFRHTWTETFEKHNFSFYPTFOSHPNSSMP----- 738  
Db 654 GGPVLVITYCYCGDHLNFLRRKAEAMHGP-----SLSPGQDSEGDSSYKNHLE 703

Qy 739 -----GSREVOIHDPDSOISLHGNPSHSEDEIEYENQKRLSEEDLNVLTFEDL 788  
Db 704 KKYVRDGSFSSQGVDTYVEMRPVSTSSDFFKQD-LDKEHSRPLE-----LWDL 753

Qy 789 LCFAYQVAKGMEFEFKSCVHRDLAARNVLVTHGKVKICDFGLARDIMSDSNVYVRGNA 848  
Db 754 LHFSSQVAGMAFLASKNCIHRDVAARNVLLTSGHVAKIGDFGLARDIMSDSNVYVRGNA 813

Qy 849 RLPVKWMAPEFEGIYTIKSDVWSYGILLWEIFSLGVNYPGIPVDANFYKLIQNGFKM 908  
Db 814 -LPVKWMAPEFICDVIYTVQSDVWSYGILLWEIFSLGLNYPGIGHVNNKFEYKLVKDGQYM 872

Qy 909 DQFYATEEYIIMQSCWAFDSKRPSPNLTSLFCQLADAEAMYQNVDCRVSECP 966  
Db 873 AQPVFAPKNYISIMQSCWDLPTRRPTFOQICFL-----QEQARLERRDQDYANLP 924





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:18:49 ; Search time 18.0091 seconds  
(without alignments)  
2286.959 Million cell updates/sec

Title: US-09-919-408-4  
Perfect score: 5274  
Sequence: 1 MPALARDAGTVPLEWFSAM.....PFSREMDLGLLSPOAQVEDS 993

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5274	100.0	993	1 FLT3_HUMAN	P36888 homo sapien
2	4429.5	84.0	992	1 FLT3_MOUSE	Q00342 mus musculus
3	1295.5	24.6	978	1 KIT_CAPHI	Q28317 capra hircu
4	1286	24.4	977	1 KIT_BOVIN	P43481 bos taurus
5	1266	24.0	975	1 KIT_MOUSE	P05532 mus musculus
6	1251	23.7	976	1 KIT_HUMAN	P10721 homo sapien
7	1247.5	23.7	975	1 KIT_CANFA	O97799 canis famil
8	1225.5	23.2	980	1 KPM5_FELCA	P13369 felis silve
9	1224.5	23.2	978	1 KIT_FELCA	Q28889 felis silve
10	1224	23.2	978	1 KPM5_FSVMD	P00545 feline sarc
11	1216.5	23.1	972	1 KPM5_HUMAN	P07333 homo sapien
12	1215	23.0	960	1 KIT_CHICK	Q08156 gallus gall
13	1204	22.8	977	1 KPM5_MOUSE	P09581 mus musculus
14	1194	22.6	978	1 KPM5_RAT	Q00495 mus musculus
15	1176	22.3	1088	1 PGDS_RAT	P20786 rattus norv
16	1169	22.2	1089	1 PGDS_MOUSE	P26618 mus musculus
17	1166	22.1	1089	1 PGDS_HUMAN	P16234 homo sapien
18	1128	21.4	1087	1 PGDS_XENLA	P26619 xenopus lae
19	1105.5	21.0	1098	1 PGDR_MOUSE	P05622 mus musculus
20	1086.5	20.6	1106	1 PGDR_HUMAN	P09619 homo sapien
21	990	18.8	1338	1 VGR1_HUMAN	P17948 h vasculat
22	989.5	18.8	370	1 KIT_FSVHZ	P04048 feline sarc
23	959.5	18.2	1333	1 VGR1_MOUSE	P35969 mus musculus
24	956	18.0	1356	1 VGR1_RAT	P53767 rattus norv
25	951.5	18.0	1356	1 VGR1_HUMAN	P35968 homo sapien
26	935	17.7	1348	1 VGR2_COTJA	P52583 coturnix co
27	923	17.5	1367	1 VGR2_MOUSE	P35918 mus musculus
28	917.5	17.4	1343	1 VGR2_RAT	O08775 rattus norv
29	907.5	17.2	1363	1 VGR3_MOUSE	P35917 mus musculus
30	904.5	17.2	1298	1 VGR3_HUMAN	P35916 homo sapien
31	874.5	16.6	823	1 CEK3_CHICK	P18461 gallus gall
32	849	16.1	821	1 FGR2_HUMAN	P21802 homo sapien
33	845	16.0	813	1 FGR2_XENLA	Q03364 xenopus lae

## ALIGNMENTS

## RESULT 1

## FLT3\_HUMAN

ID FLT3\_HUMAN STANDARD; PRT; 993 AA.

AC P36888; Q13414;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE FL cytokine receptor precursor (EC 2.7.1.112) (Tyrosine-protein kinase receptor FLT3) (Stem cell tyrosine kinase 1) (STK-1) (CD135 antigen).

GN FLT3 OR STK1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID:9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-94119906; PubMed-7507245;

RA Small D., Levenstein M., Kim E., Carow C., Amin S., Rockwell P.,

RA Witte L., Burrow C., Ratajczak M.Z., Gewirtz A.M., Civin C.I.;

RT "STK-1, the human homolog of Flk-2/Flt-3, is selectively expressed in

RT CD34+ human bone marrow cells and is involved in the proliferation of

RT early progenitor/stem cells.";

RL Proc. Natl. Acad. Sci. U.S.A. 91:459-463(1994).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE-93357464; PubMed-8394751;

RA Rosnet O., Schiff C., Pebusque M.J., Marchetto S., Tonnellet C.,

RA Toiron Y., Birg F., Birnbaum D.;

RT "Human FLT3/FLK2 gene: cDNA cloning and expression in hematopoietic

RT cells.";

RL Blood 82:1110-1119(1993).

RN [3]

RP SEQUENCE OF 783-942 FROM N.A.

RC TISSUE=Testis;

RX MEDLINE-91169547; PubMed-2004790;

RA Rosnet O., Mattei M.-G., Marchetto S., Birnbaum D.;

RT "Isolation and chromosomal localization of a novel FMS-like tyrosine

RT kinase gene.";

RL Genomics 9:380-385(1991).

CC -!- FUNCTION: RECEPTOR FOR THE FL CYTOKINE. HAS A TYROSINE-PROTEIN

CC KINASE ACTIVITY.

CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein

CC tyrosine phosphate.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: BONE MARROW CELLS.

CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-

CC PROTEIN KINASES.

CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD135 entry;

CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd135.htm"

CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;

CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/FLT3ID144.html"

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: 002687; AAA18947.1; -  
CC EMBL: 226652; CAA81393.1; -  
CC EMBL: L36162; AAA35487.1; -  
CC HSSP: P11362; 1FGK.  
CC Genew: HGNC:3765; FLT3.  
CC MIM: 136351; -  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003600; Ig\_like.  
DR InterPro: IPR001824; RTK\_kinaseIII.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR Pfam: PF00047; Ig; 1.  
DR Pfam: PF00069; pkinase; 1.  
DR ProDom: PD000001; Euk\_pkinase; 2.  
DR SMART: SM00410; IG\_like; 1.  
DR SMART: SM00219; Tyr\_KC; 1.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE: PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR Signal: Transferrase; Tyrosine-protein kinase; Receptor; Transmembrane;  
KW Glycoprotein; Phosphorylation; ATP-binding; Immunoglobulin domain.  
FT SIGNAL 1  
FT CHAIN 27 993  
FT DOMAIN 27 543  
FT TRANSMEM 544 563  
FT DOMAIN 564 993  
FT DOMAIN 610 943  
FT NP\_BIND 616 624  
FT BINDING 644 644  
FT ACT\_SITE 811 811  
FT CARBOHYD 43 43  
FT CARBOHYD 100 100  
FT CARBOHYD 151 151  
FT CARBOHYD 306 306  
FT CARBOHYD 323 323  
FT CARBOHYD 351 351  
FT CARBOHYD 354 354  
FT CARBOHYD 473 473  
FT CARBOHYD 502 502  
FT CARBOHYD 541 541  
FT CONFLICT 8 8  
FT CONFLICT 10 11  
FT CONFLICT 78 78  
FT CONFLICT 227 227  
FT CONFLICT 346 346  
FT CONFLICT 940 940  
SQ SEQUENCE 993 AA; 112804 MW; 16790124B02F6BBF CRC64;  
  
Query Match 100.0%; Score 5274; DB 1; Length 993;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 993; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MPALARDAGTVLLVVFSAAMIFGTITNQDLPIVKVLINHKNDSSVGVKSSSYPMVSESP 60  
Db 1 MPALARDAGTVLLVVFSAAMIFGTITNQDLPIVKVLINHKNDSSVGVKSSSYPMVSESP 60  
  
Qy 61 EDLGCALRQSSGTVYAAAEVDVDSASITLQVLVDAPGNISCLWFKHSSLNACQPHFDL 120  
Db 61 EDLGCALRQSSGTVYAAAEVDVDSASITLQVLVDAPGNISCLWFKHSSLNACQPHFDL 120  
  
Qy 121 ONRGVSVWILKMTQAGEYLLFIQSEATNYTLFTVSRINTLLYTLRPRYFRKMNQD 180  
Db 121 ONRGVSVWILKMTQAGEYLLFIQSEATNYTLFTVSRINTLLYTLRPRYFRKMNQD 180  
  
Qy 181 ALVCISSEVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLFGDTRCCARNELGRE 240  
Db 181 ALVCISSEVPEPIVWVLCDSQGESCKEESPAVVKKEKVLHFLFGDTRCCARNELGRE 240

Qy 241 CTRLFITDLNQTPTTLPOLFLKVGEPILWIRKAVHVNHGFLTWELNKALEEGNYFEM 300  
Db 241 CTRLFITDLNQTPTTLPOLFLKVGEPILWIRKAVHVNHGFLTWELNKALEEGNYFEM 300  
  
Qy 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEYE 360  
Db 301 STYSTNRTMIRILFAFVSSVARNDTGYTCSSSKHPSQSALVTIVGKGFINATNSSEYE 360  
  
Qy 361 IDOYEEFCFSVRFKAYPOIRCTWTFSRKSFCEQKGLDNGYSISKFCNKHKHPGEYIHA 420  
Db 361 IDOYEEFCFSVRFKAYPOIRCTWTFSRKSFCEQKGLDNGYSISKFCNKHKHPGEYIHA 420  
  
Qy 421 ENDDAQFTKMETLNIARRKPOVLAEASASOACFSOGYPLPSMTWKCKSDKSPNCTEEITE 480  
Db 421 ENDDAQFTKMETLNIARRKPOVLAEASASOACFSOGYPLPSMTWKCKSDKSPNCTEEITE 480  
  
Qy 481 GWNRKANKRVFGQWVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQD 540  
Db 481 GWNRKANKRVFGQWVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSPGPPFIQD 540  
  
Qy 541 NISFYATIGVCLLFTIWLTLICHKKQFVYESOLQWVQVVTGSSDNEYFYVDFREYED 600  
Db 541 NISFYATIGVCLLFTIWLTLICHKKQFVYESOLQWVQVVTGSSDNEYFYVDFREYED 600  
  
Qy 601 LKWEPPRENLEFGKVLGSCAFKVMNATAYGISKTCVSIQVAVKMLKEKADSSERREALMS 660  
Db 601 LKWEPPRENLEFGKVLGSCAFKVMNATAYGISKTCVSIQVAVKMLKEKADSSERREALMS 660  
  
Qy 661 ELKMMTQLGSHENIVNLGACTLSGPIYLIFCYCCYDGLLNLRSRREKFRHTWTTEIFE 720  
Db 661 ELKMMTQLGSHENIVNLGACTLSGPIYLIFCYCCYDGLLNLRSRREKFRHTWTTEIFE 720  
  
Qy 721 HNFSEFYPTQSHPNSSMPGSRVQIHPDSDQISGLHGNFSHSEDEIYEYENQKLEEEEDL 780  
Db 721 HNFSEFYPTQSHPNSSMPGSRVQIHPDSDQISGLHGNFSHSEDEIYEYENQKLEEEEDL 780  
  
Qy 781 NVLTPEDLCLCFAYQVAKGMEFFLEKSCVHRDLAARNVLTGHGVKVKICDFGLARDMSDS 840  
Db 781 NVLTPEDLCLCFAYQVAKGMEFFLEKSCVHRDLAARNVLTGHGVKVKICDFGLARDMSDS 840  
  
Qy 841 NYVVRGNARLPVKWMAPELSFEGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900  
Db 841 NYVVRGNARLPVKWMAPELSFEGIYTIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900  
  
Qy 901 LIQNGFKMDQPPYATETEEIYIIMOSWAFDSRKRSPFNLTSLFCGLQADAEAEAMYNQVNDG 960  
Db 901 LIQNGFKMDQPPYATETEEIYIIMOSWAFDSRKRSPFNLTSLFCGLQADAEAEAMYNQVNDG 960  
  
Qy 961 RVSECPHTYQNRPRFSPREMDLGLLSPOAQVEDS 993  
Db 961 RVSECPHTYQNRPRFSPREMDLGLLSPOAQVEDS 993  
  
RESULT 2  
FLT3\_MOUSE STANDARD; PRT; 992 AA.  
AC Q00342;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
DE FL cytokine receptor precursor (EC 2.7.1.112) (Tyrosine-protein kinase receptor flk-2) (Fetal liver kinase 2) (Tyrosine-protein kinase FLT3).  
GN FLT3 OR FLK-3 OR FLK-2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=91292518; PubMed=1648448;  
RA Matthews W., Jordan C.T., Wiegand G.W., Pardoll D., Lemischka I.R.;

RT	"A receptor tyrosine kinase specific to hematopoietic stem and progenitor cell-enriched populations.";
RL	Cell 65:1143-1152(1991).
RP	[2]
RN	SEQUENCE FROM N.A.
RX	MEDLINE=92019834; PubMed=1656368;
RA	Rosnet O., Marchetto S., Delapeyriere O., Birnbaum D.;
RA	"Murine Flt3, a gene encoding a novel tyrosine kinase receptor of the PDGFR/CSF1R family.";
RL	Oncogene 6:1641-1650(1991).
RN	[3]
RP	CHARACTERIZATION.
RX	MEDLINE=93205405; PubMed=8384358;
RA	Maroc N., Rotapel R., Rosnet O., Marchetto S., Lavezzi C.,
RA	Manno P., Birnbaum D., Dubreuil P.;
RT	"Biochemical characterization and analysis of the transforming potential of the FLT3/FLK2 receptor tyrosine kinase.";
RL	Oncogene 8:909-918(1993).
CC	-1- FUNCTION: RECEPTOR FOR THE FL CYTOKINE. HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC	-1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.
CC	-1- TISSUE SPECIFICITY: HEMATOPOIETIC STEM AND PROGENITOR CELL-ENRICHED POPULATIONS. FOUND IN BRAIN, PLACENTA AND TESTIS.
CC	-1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-PROTEIN KINASES.
CC	-1- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcements">http://www.isb-sib.ch/announcements</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
EMBL	M64689; AAA37634.1; -
DR	EMBL; X59398; CAA42041.1; -
DR	PIR; A39931; A39931.
DR	HSP; P11362; IFGK.
DR	MGI; MGI:95559; Flt3.
DR	InterPro; IPR000719; Euk_pkinase.
DR	InterPro; IPR003006; Ig_MHC.
DR	InterPro; IPR003600; Ig_like.
DR	InterPro; IPR001824; RTKinaseII.
DR	InterPro; IPR001245; Tyr_pkinase.
DR	Pfam; PF00047; Ig_1.
DR	Pfam; PF00069; pkinase; 1.
DR	ProDom; PD000001; Euk_pkinase; 2.
DR	SMART; SM00410; IG_like; 1.
DR	SMART; SM00219; TyrKc; 1.
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR	PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR	Signal; Transference; Tyrosine-protein kinase; Receptor; Transmembrane; Glycoprotein; Phosphorylation; ATP-binding; Immunoglobulin domain.
KW	FT SIGNAL 1 27 POTENTIAL.
KW	FT CHAIN 28 932 FL CYTOKINE RECEPTOR.
FT	FT DOMAIN 28 544 EXTRACELLULAR (POTENTIAL).
FT	FT TRANSMEM 545 564 POTENTIAL.
FT	FT DOMAIN 565 992 CYTOPLASMIC (POTENTIAL).
FT	FT DOMAIN 611 946 PROTEIN KINASE.
FT	FT NP_BIND 617 625 ATP (BY SIMILARITY).
FT	FT BINDING 645 645 ATP (BY SIMILARITY).
FT	FT ACT_SITE 814 814 BY SIMILARITY.
FT	FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

Db 841 SDSSVVRGNARLPKWNAPESLFGIYTIKSDVWSYGLLLWEIFSLGVNPPYGPVDAN 900
Qy 898 FYKLIONGFKMDOPYATEEIIYIMOSWAFDSRKRPPNLTSLFGCOLAAEAMYNQ 957
Db 901 FYKLQSGFKMEQPEYATEGIIYVMSWAFDSRKRPPNLTSLFGCOLAAEAEAC 957
Qy 958 VDGVRSECPHYQNRPPSRENDLGLLSPQAQVE 991
Db 958 ----IRTSIHLPKQAAPQORG-GLRAQSPORQVK 986

RESULT 3
KIT_CAPHI
ID KIT_CAPHI STANDARD; PRT; 978 AA.
AC Q28317;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).
GN KIT.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_taxid=9925;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Shiba; TISSUE=Cerebellum;
RX MEDLINE=97342548; PubMed=9199245;
RA Tanaka S., Yanagisawa N., Tojo H., Kim Y.-J., Tsujimura T.,
RA Kitamura Y., Sawasaki T., Tachi C.;
RT "Molecular cloning of cDNA encoding the c-kit receptor of Shiba goats
RT and a novel alanine insertion specific to goats and sheep in the
RT kinase insert region.";
RL Biochim. Biophys. Acta 1352:151-155(1997).
CC !- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC !- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC !- SUBCELLULAR LOCATION: Type I membrane protein.
CC !- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC !- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL; D45168; BAA08116.1; -
DR HSSP; P11362; IFGK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003598; Ig_C2.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR001824; RTKinaseIII.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 2.
DR SMART; SM00410; IG_Like; 2.
DR SMART; SM00408; IGc2; 1.
DR SMART; SM00219; TyrcK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

```

```

DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Proto-oncogene: Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW Immunoglobulin domain.
FT SIGNAL 1 22
FT CHAIN 23 978
FT DOMAIN 23 521
FT TRANSMEM 522 544
FT DOMAIN 545 978
FT DOMAIN 590 939
FT NP_BIND 596 604
FT BINDING 624 624
FT ACT_SITE 794 794
FT MOD_RES 825 825
FT CARBOHYD 94 94
FT CARBOHYD 130 130
FT CARBOHYD 145 145
FT CARBOHYD 284 284
FT CARBOHYD 294 294
FT CARBOHYD 301 301
FT CARBOHYD 321 321
FT CARBOHYD 353 353
FT CARBOHYD 368 368
FT CARBOHYD 401 401
FT CARBOHYD 464 464
FT CARBOHYD 487 487
FT SEQUENCE 978 AA; 109722 MW; CA4D663F98205CA9 CRC64;

Query Match 24.6%; Score 1295.5; DB 1; Length 978;
Best Local Similarity 33.3%; Pred. No. 3.9e-79;
Matches 335; Conservative 169; Mismatches 339; Indels 163; Gaps 35;

QY 47 VGKSSYPVMSPELDGC-ALRPQSGTVYEAANEVDVSAITQLVLDAGNISCILW 105
Db 20 VQTGSSQPSV--SPGELSLPSIHPAKS-----ELIVSVDGEIRLLCTDGFVK--W 66
QY 106 VFKHSSLNCOPHDLQNRGVSMVILKMTQTQAGEYLLFFIQSEAT---NVTILFTVSIRN 162
Db 67 TFE-----ILGOLSEKTNPEWIT-EKAENATNGTCTNKGGLSS 105
QY 163 TL-----LYTLRRPYFRKMNQDALVCISSEVPEPIVWVLCDSQGESCKEESPAV 213
Db 106 SIIVFVRDPEKFLIDLPLXGKEEN-DTLVRCPLTDEP-VTNSYSLTGCCKPLPKDLTFV 163
QY 214 -----VKKEBKVLHELFGTDIRC--CARNELGRE-CTRLETFD----LNQTPQTLTP 258
Db 164 ADPKAGITIRNVKREYHRL-----CLHCSANQKGSMLSKKFTLKVRAAIKAVPVVSVS 217
QY 259 Q--LFLKVGEPLWIRCKAVHVNHGFGLTWELENKALEEGNYFEMSTYSTNRTMIRILFAF 316
Db 218 KTSYLLREGEEFAVTCILKDVSSVDSMMWIKENSQQSKAQTCKNSWHQGFSLYRQERLT 277
QY 317 VSSVARNDTGYTCSSSKHPQSALVT---IVGKGFTNA-----TNSSEYDEIDQY 364
Db 278 ISSARVNDSGVFCYANNNTGSGSANVTTLLEVVDKGFNIEFPMNTTVFVNDGENVDL--- 334
QY 365 EEFCSVRFKAYPO-----IRCTWTFSRKSPFCQKGLDNGYSKSCFNKHKH----- 411
Db 335 -----VVEYAYPKPEHQRWIYMNRTSTDKWDYPKSE-----NESIRYVNNELHLRLK 384
QY 412 --QPGEYIFHAENDDAQFTKMFILNIRRKPOVLAEASA--SQASCFSDGSDGYPSPWKKK 467
Db 385 GTEGGTYTFHVSNDVNSSTFNVNNTKPEILTHDLVNGMLQCVAAAGPEPTIDWYFC 444
QY 468 SDKSPNCTEEI-TEGVNRRKANRKFQGWVSSSTLNNSAEIKGFLVKCCAYNSLGTSCET 526
Db 445 PGTEQRCSVPVGPVDVQIQNSVSPFCKLVVYSTIDDTSTFKHNGTVECRAYNDVGKSSAS 504
QY 527 ILLNSPGPPFFIODN-----ISFYATIGCVCLLFIVVLTLICHKYKKFY 572
Db 505 F-----NFAFKGNKEQIHAHTLFTPLLLIGFVIAAGLMCIFVMILT----YKYLQKPMY 554

```

QY 573 ESQOMV-QVTGSSDNEYFYDFREYEDYDKWFFPRENLEFGKVLGSGAFGKVMNATAYG 631  
 Db 555 EVQWKVVEING---NNVYIDPTOLPYDHKWEFPNRLSFKTGLGAGAGKVVETAYG 611  
 QY 632 ISKTGVSQVAVKMLKEKADSSERALMSLKMWTOLGSHENIVNLLGACTISGPIYLIF 691  
 Db 612 LTKSDAAMTVAVKMLKPSAHLTERALMSLKVSLYGLNHNMINVLLGACTIGGPTLVIT 671  
 QY 692 EYCCYGDLLNLYRSKREKFRHTWTE-----IFKEHNEFYPTFQSHPNSSM---PGSRE 742  
 Db 672 EYCCYGDLLNLYRRKRDSEFICKQEDHAEVALYKNLLHKSCKSSNDSTNEYMDMPGVSY 731  
 QY 743 VOIHPSDQISLHGNHSEHSEIEYENQKRLSEEDNLVLFEDLCLFAYQVAKGMEFL 802  
 Db 732 VVPTKAADKRRSARIGSY-----IERDVTPTAMEDDEL-ALDLEDLSFSYQVAKGMFL 785  
 QY 803 EFKSCVHRDLAARNVLYHGKVVVKICDGLARDIMSDSNVYVGRNARLPVKWAPESLFE 862  
 Db 786 ASKNCIHRDLAARNLLTHGRITKICDFGLARDIKNSNVYVGNARLPVKWAPESIFN 845  
 QY 863 GIYTIKSDWSYGILLWEIFSLGVNYPGIPVDANFYKLIQNGKMDQPEYATEEYIIM 922  
 Db 846 CVYTESDWSYGIFLWELFSLGSPYPCGMPVDSFYKMKIEGFRMLSPHAPAEYIM 905  
 QY 923 OSCWAFDSKRKPSFPLNLSFLGCLADAEAMYQNVGRVSEC-PH 967  
 Db 906 KTCWDADPLKRTFTFOIVOLIEKQISESTNHIYSN----LANCSPH 947  
 RESULT 4  
 KIT\_BOVIN  
 ID KIT\_BOVIN STANDARD; PRT; 977 AA.  
 AC P43481;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Mast/stem cell growth factor precursor (EC 2.7.1.112) (SCFR)  
 DE (Proto-oncogene tyrosine-protein kinase kit) (c-kit).  
 GN KIT.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 ON NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94215924; PubMed=7512939;  
 RA Kubota T., Hikono H., Sasaki E., Sakurai M.;  
 RT "Sequence of a bovine c-kit proto-oncogene cDNA."  
 RL Gene 141:305-306(1994).  
 CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL  
 CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.  
 CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND  
 CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL  
 CC 3-KINASE (PI3K).  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; D16680; BAA04084.1;  
 DR HSSP; P11362; IFK.  
 DR InterPro; IPR000719; Euk\_pkinase.

InterPro; IPR003006; Ig\_MHC.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR003600; Ig\_like.  
 DR InterPro; IPR01824; RtkinaseIII.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00047; Ig; 1.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 2.  
 DR SMART; SM00410; Ig\_like; 2.  
 DR SMART; SM00408; Ig\_c2; 1.  
 DR SMART; SM00219; Tyrc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 KW Proto-oncogene; tyrosine-protein kinase; Receptor; Transmembrane;  
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;  
 KW Immunoglobulin domain.  
 FT SIGNAL 1 22  
 FT CHAIN 23 977  
 FT DOMAIN 23 521  
 FT TRANSMEM 522 544  
 FT DOMAIN 545 977  
 FT DOMAIN 590 938  
 FT NP\_BIND 596 604  
 FT BINDING 624 624  
 FT ACT\_SITE 793 793  
 FT MOD\_RES 824 824  
 FT CARBOHYD 94 94  
 FT CARBOHYD 130 130  
 FT CARBOHYD 145 145  
 FT CARBOHYD 284 284  
 FT CARBOHYD 284 294  
 FT CARBOHYD 301 301  
 FT CARBOHYD 321 321  
 FT CARBOHYD 353 353  
 FT CARBOHYD 368 368  
 FT CARBOHYD 401 401  
 FT CARBOHYD 454 454  
 FT CARBOHYD 487 487  
 SQ SEQUENCE 977 AA; 109685 MW; 4B2719050883B7EF CRC64;  
 Query Match 24.4%; Score 1286; DB 1; Length 977;  
 Best Local Similarity 33.2%; Pred No 1.7e-78;  
 Matches 332; Conservative 172; Mismatches 344; Indels 152; Gaps 36;  
 QY 47 VQKSSYPMVSESPDLGC-ALRPQSSGYVEAAAVEVDVSASITLQVLVDAPGNISCLW 105  
 Db 20 VQTGSSQPSV--SPGELSLPSIHPAKS-----ELIVSGVDEIRLLCTDPGFVK--W 66  
 QY 106 VFKHSLNCQPHFDLONRGVSVVILKMTETQAGEYLLFIQSEAT---NYTILFTVSIRN 162  
 Db 67 TFE-----ILQLSEKTNPEWIT-EKAEATNTGNYTCTNKGGLSS 105  
 QY 163 TL-----LYTLRRPYFRKMNODALVCISESPEPIVENVLDCDSOGESKEESPAY 213  
 Db 106 SIYFVRDPEKFLDLIDPLYGKEEN-DTLVRCPLTDPPE-VTNYSLTGCEGKPLPKDLTFV 163  
 QY 214 -----VKKEEKLVELFGLTGDIRC--CARNELGRE-CTRLEFTID----LNQTPQTLTP 258  
 Db 164 ADPKAGITIRNVKREYHRL-----CLHCSANQSGKMSLSKFTLVKRAIKAVPVSVS 217  
 QY 259 Q--LFLKVGEPWIRKAVVHNVHFGLTWELENKALEEGNYFEMSTYSTNRTMIRILFAF 316  
 Db 218 KTSYLLRREGFAVTCIKIDVSSVDSMMWIKENSQOTKAOTKKNSWHQGFVSLYRLERT 277  
 QY 317 VSSVARNDTGYTCSSSKHPSQSALVT---IVCKGGINA-----TNSSDEYEDIQY 364  
 Db 278 ISSARVNDSGVFCMYANNTFGSANVTTLLEVVDKGINFIPPMNTTVFVNDGENVDL--- 334  
 QY 365 EEFECFSVREKAYPO-IRCTWTTSRKSS-----FPCQKGLDNGYSISKFNHKK----- 411  
 Db 335 -----VVEYEATPKPVHROWIYANRTSTDKWDDYPKSE-----NESNIRYVNEHLHLTRK 384

```

QY 412 --QCEYIFHAENDDAQTKMFTNIRKPOVLAEASA--SOASCFSDCGYPLPSTWTKC 467
Db 385 GTEGTYTFHVSNSVNSVFNVTNKPEILTHDLVNGMLQCVAGFPEPTIDWFC 444
QY 468 SDKSPNCTEEI-TEGVWNRKANRKYVGQWSSSTLMSAIFGLFKVCAKYNLSGTSCT 526
Db 445 PGEQRCSVPVGPVDVQIONSSVSPGKLKLVYSTDDSTFKNGTVCEARAYNDVGKSSAS 504
QY 527 ILLNSPGFPP-----FIQDNISFYATIGVCLFIVVLITLLCHKKYKQFRYSQLOM 578
Db 505 FNFAFGKNSKEQIHAHTLFTPLLLGFIAGLMCFVMTL----YKYLQKPMYEVQWKV 560
QY 579 V-OVTGSSDNEYVDFEYEDLQWEPRENLEFGVLGSAFGKVMNATAYGSKTCV 637
Db 561 VEEING-----NNVYIDPQLPDHKEWEPNRLSFGKLGAGAFKGVETAYGLKSDA 617
QY 638 SIQVAVKMLKEKADSSEREAALMSLKMOTLGSHENIVNLLGACTLSGPIYLIFEYCCYG 697
Db 618 AMTAVAVKMLKPSAHLTEREALMSLKVLSYLGHNHNVNLLGACTIGGPTLVITEYCCYG 677
QY 698 DLLNVLRSKREKFTWTE-----IFKEHNSFYPTFQSHPNSSM---PGSREVOIHFD 748
Db 678 DLLNVLRRKRDSCQDHAELVAALYKLLHRSKSSCNDSTNEYMDMKPGVSYV-VPFK 736
QY 749 SPOISGLHNSPHSEDEIEYENQKLEEREEDLVLTFFEDLLCFAYQVAKGMFELEFKSCV 808
Db 737 ADKRSARIGSV-----IERDVTPEIMDEDEL-ALDELLSFSQVAKGMFLASKNCI 790
QY 809 HRDLAARNVLTHGVKVKICDFGLARDIMSDSNVYVGRNARLPVKWMAPESLFEGIYTIK 868
Db 791 HRDLAARNILLTHGRITKICDFGLARDIKDSDNVYVGRNARLPVKWMAPESFNCVYTFE 850
QY 869 SDVWSYGILLWEIFSLGVNVPYGPVDFANFYKIQNGFKMDOPPFATIEIYIMQSCWAP 928
Db 851 SDVWSYGIFLWELFSLGSSPYGPMVDKSFYKMKRFGFRLSPHAPAEAMYDIMKTCWDA 910
QY 929 DSRKRPSPNLTSLFLGCOLADAEAEAMQVNDGRVSPC-PH 967
Db 911 DPLKRPTRQIVOLIEKQISESTNHIYSN-----LANCSPH 946

RESULT 5
KIT_MOUSE
ID KIT_MOUSE STANDARD: PRT: 975 AA.
AC P05332; Q61415; Q61416; Q61417;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
GN KIT OR SL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A..
RC STRAIN=BALB/C;
RX MEDLINE=88296403; PubMed=2456920;
RA Qiu F., Ray P., Brown K., Barker P.E., Jhanwar S., Ruddle F.H.,
RA Besmer P.;
RT "Primary structure of c-kit: relationship with the CSF-1/PDGF
RT receptor kinase family -- oncogenic activation of v-kit involves
RT deletion of extracellular domain and C terminus.";
RL EMBO J. 7:1003-1011(1988).
RN [2]
RP SEQUENCE FROM N.A. (TRUNCATED FORM).
RC STRAIN=ICR;
RX MEDLINE=92331813; PubMed=1378413;
RA Rossi P., Marziani G., Albanesi C., Charlesworth A., Geremia R.,
RA Sorrentino V.;
RT "A novel c-kit transcript, potentially encoding a truncated receptor,
```

```

RT originates within a kit gene intron in mouse spermatids.";
RL Dev. Biol. 152:203-207(1992).
RN [3]
RP LIGAND.
RX MEDLINE=91006023; PubMed=1698611;
RA Tan J.C., Buck J., Levi E., Besmer P.;
RT "Candidate ligand for the c-kit transmembrane kinase receptor: KL, a
RT fibroblast derived growth factor stimulates mast cells and erythroid
RT progenitors.";
RL EMBO J. 9:3287-3294(1990).
RN [4]
RP VARIANT W42 ASN-790.
RX MEDLINE=90100577; PubMed=1688471;
RA Tan J.C., Nocka K., Ray P., Traktman P., Besmer P.;
RT "The dominant W42 spotting phenotype results from a missense mutation
RT in the c-kit receptor kinase.";
RL Science 247:209-212(1990).
RN [5]
RP VARIANTS W37 LYS-582; WV MET-660 AND W41 MET-831.
RX MEDLINE=90269214; PubMed=1693331;
RA Nocka K., Tan J.C., Chiu E., Chu T.Y., Ray P., Traktman P.,
RA Besmer P.;
RT "Molecular bases of dominant negative and loss of function mutations
RT at the murine c-kit/white spotting locus: W37, WV, W41 and W1.";
RL EMBO J. 9:1805-1813(1990).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DISEASE: WHITE-SPOTTING VARIANTS INDUCES SEVERE EFFECTS ON
CC PIGMENTATION, GAMETOGENESIS AND HEMATOPOIESIS. MICE HOMOLOGOUS
CC FOR W42 DIE PERINATALLY OF MACROCYTIC ANEMIA.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y00864; CAA68772.1; -
CC EMBL; X65997; CAA46798.1; -
CC EMBL; X65998; CAA46799.1; ALT_SEQ.
CC EMBL; X65998; CAA46800.1; -
CC PIR; S00474; TWSKRT.
CC HSSP; P11362; IFGK.
CC MGD; MGI:966677; Kit.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003598; Ig_C2.
CC InterPro: IPR003600; Ig_Like.
CC InterPro: IPR001824; RTKinaseIII.
CC InterPro: IPR001245; Tyr_pkinase.
CC Pfam; PF00047; ig; 1.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 2.
CC SMART; SM00410; IG_Like; 3.
CC SMART; SM00408; IGC2; 1.
CC SMART; SM00219; Tyrc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
CC Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
KW
```

Immunoglobulin domain; Disease mutation.

KW SIGNAL 1 22  
FT CHAIN 23 975 MAST/STEM CELL GROWTH FACTOR RECEPTOR.  
FT DOMAIN 23 519 EXTRACELLULAR (POTENTIAL).  
FT TRANSEM 520 542 POTENTIAL.  
FT DOMAIN 543 975 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 588 935 PROTEIN KINASE.  
FT NP\_BIND 594 602 ATP (BY SIMILARITY).  
FT BINDING 622 622 ATP (BY SIMILARITY).  
FT ACT\_SITE 790 790 BY SIMILARITY.  
FT MOD\_RES 821 821 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT CARBOHYD 146 146 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 296 296 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 303 303 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 323 323 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 355 355 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 370 370 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 466 466 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 489 489 N-LINKED (GLCNAC...) (POTENTIAL).  
FT VARIANT 1 12 MRGARGAMDLLC -> MAVAVFPFLPQQ (IN TRUNCATED FORM).  
FT VARIANT 13 785 MISSING (IN TRUNCATED FORM).  
FT VARIANT 582 582 E -> K (IN W37 SPOTTING).  
FT VARIANT 660 660 T -> M (IN W4 SPOTTING).  
FT VARIANT 790 790 D -> N (IN W42 SPOTTING).  
FT VARIANT 831 831 V -> M (IN W41 SPOTTING).  
SQ SEQUENCE 975 AA; 109001 MW; BABSCA4D9AF9CD2A CRC64;

Query Match 24.0%; Score 1266; DB 1; Length 975;  
Best Local Similarity 33.2%; Pred. No. 3.7e-77;  
Matches 336; Conservative 173; Mismatches 330; Indels 172; Gaps 41;

QY 35 CVLINHKNDSSVCKSSYPVMSPEPDLGCLALRQSSCTVYEAARVVDVSASITLQVL 94  
DB 12 CVLLVLLRGQATSPASPEPSP-----SIHPAQSS-----ELIVEAGDTL----- 54  
QY 95 VDAPGNISCL-----WYFKHSSLNCPHFQDLQNGVSVMLKMTQTQAGYLLFIOSE 148  
DB 55 -----SLTCDIDPDFVRWTFK-----IYFN-----ENVENKKNWQI-EKAE 89  
QY 149 ATNVTILETVSIRNPL-----LYTLRPYFRKMNQDALVCISSEVPPIVE 195  
DB 90 ATR-TGTYTCSNGLTSIIYVFRDPAKFLVGLPLFGK-EDSDALVRCPLTDPQ-VSN 146  
QY 196 WVLCDSCQESCKEESPAV-----VKKEKVLHELCTDIRCCARNELGRECTRL--- 244  
DB 147 YSLIECDGKSLPTDLTFVFNPKAGTIRKVRKAYRLC---VRCQAQ-----RDGTWLHSD 199  
QY 245 -FTIDLNOT---POTTLPLQ--FLKVGEPILWIRCKAVHVNHGFLTW-----E 286  
DB 200 KFTLVKRAIKAIPIVSVPEPESHLLKKGDTFVVCITKDVSTSVNSMWLKNMPQPHIAQ 259  
QY 287 LENKALEGNYFEMSTYNTNTRMIRILFAFVSSVARNDGTYYTSSSKHPSQALVT--- 343  
DB 260 VKHNSWHRGDF-----NYERQETLT-----ISSARVDSGVFMCYANNTFGSANVTTLK 309  
QY 344 IVGKGFINATN-SSEDIYIDQYEEFCFSVRKAYPQ-IRCTWTFPSKFPQCKGLD--- 398  
DB 310 VVEKGFINISPVKNTVFVTGDNVLDVVEYEAYPEKPEHQOYIYNNRT--SANKGKDYVK 367  
QY 399 -NGYSISKFCNKH-----OPGEYIHAENDDAQFKMETLINIRRPQVLA--EASAS 448  
DB 368 SDNKSNIYVQLRLTRUKGTEGGTYTFLVNSDSASAVTFNVVNTPEILTYDRLING 427  
QY 449 QASCFSDGYPPLPFWTKKCDKSPNCTEITE-GYWNKANKKVFQGVSSSTLNKSEAI 507  
DB 428 MLCQVAEGFPEPTIDWYECTGAEQRCITPVPSPDVQVQVNSVSPFGKLVQSSIDSSVFR 487  
QY 508 KGLVKCAYNSLGTSCETILLNSPGPPFF-----IODN-----ISFYATIGVCILLFIW 557  
DB 488 HNGTVECKASNDVGKS--SAFFN-----FAFREIQIAHTLFTPLLCGFVVAACAMGLIWMV 541  
QY 558 LTLILCHYKQKQFRYESQLQMV-QVTGSSDNEYFYVDREYEDYLUKWEFFRENLEFGKVL 616

Db 542 LT-----YKYLQKPMYEVQKVVVEEING---NNVYIDPTQLPYDHHKWEPRNLSFGKTL 594  
QY 617 GSGAFGKVNATAYGISKTVGSIQVAVKMLKEKADSSEREAALMSELKMMTQLOSHENIVN 676  
Db 595 GAGAFGKVVYATAYGLIKSDAAMTAVAKMLKPSAHLTREALMSELKVLKSYLGNHMINVN 654  
QY 677 LLGACTLSGPVILIPFECYCGDLLNLYLRSKREF-----HRTWTEIFKEHNEFSYPTFQ 730  
Db 655 LLGACTVGGTTLVITEYCCYCGDLLNLFRRKRDSFIKSOEEQAEALYKLLHSTPEFSD 714  
QY 731 SHPNSSM---PGSREVIQIHPDSQDISLHGNSPHSEDEIEYENQKLEEDLNVLTFED 787  
Db 715 S-SNEYMDMKPGSVV-VPTKTDKRSARDSY-----IERDVTPALMEDEL-ALDLD 766  
QY 788 LLCFAVQVAKGMEFFLEFKSCVHRDLAARNVLYTHGKVKICDFGLARDIMSDSNVYVGRN 847  
Db 767 LLSFSYQVAKAMAFASKNCIHRDLAARNLLTHGRITKICDFGLARDIRNDSNVYVWGN 826  
QY 848 ARLPVKWMAPESLFEGYTIKSDVMSYGGILLAEIFSLGVNYPGIPVDANFYKLIQNGFK 907  
Db 827 ARLPVKWMAPESLFSCVYTFESDVMSYGFILWELFSLGSSPYGMPVDSFKYMKIEGFR 886  
QY 908 MDQPFYATEEIIYIMOSCWAFDSKRKPSFNLTSLGCOLADAEAMYQV 958  
Db 887 MYSPEHAPAEYMDVMTCDADPLKRPTEKQVQLIEKQISDSTKHIYSNL 937

RESULT 6  
KIT\_HUMAN  
ID KIT\_HUMAN STANDARD; PRT; 976 AA.  
AC P10721.  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 01-JUL-1989 (Rel. 11, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Mast/stem cell growth factor precursor (EC 2.7.1.112) (SCFR)  
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit) (CD117 antigen).  
GN KIT.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=term placenta, and Fetal brain;  
RX MEDLINE=88111521; PubMed=2448137;  
RA Yarden Y., Kuang W.-J., Yang-Feng T., Coussens L., Mumentz S.,  
RA Dull T.J., Chen E., Schlessinger J., Francke U., Ullrich A.;  
RT "Human proto-oncogene c-kit: a new cell surface receptor tyrosine  
RT kinase for an unidentified ligand.";  
RL EMBO J. 6:3341-3351(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93064697; PubMed=1279499;  
RA Giebel L.B., Strunk K.M., Holmes S.A., Spritz R.A.;  
RT "Organization and nucleotide sequence of the human KIT (mast/stem  
RT cell growth factor receptor) proto-oncogene.";  
RL Oncogene 7:2207-2217(1992).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97179223; PubMed=9027509;  
RA Andre C., Hampe A., Lachaume P., Martin E., Wang X.P., Manus V.,  
RA Hu W.X., Galibert F.;  
RT "Sequence analysis of two genomic regions containing the KIT and the  
RT FMS receptor tyrosine kinase genes.";  
RL Genomics 39:216-226(1997).  
RN [4]  
RP VARIANT LYS-583.  
RX MEDLINE=92291284; PubMed=1376329;  
RA Fleischman R.A.;  
RT "Human piebald trait resulting from a dominant negative mutant allele  
RT of the c-kit membrane receptor gene.";  
RL J. Clin. Invest. 89:1713-1717(1992).

[5] VARIANT LEU-584.  
 CC MEDLINE-92133600; PubMed-1370874;  
 RA Spritz R.A., Giebel L.B., Holmes S.A.;  
 RT "Dominant negative and loss of function mutations of the c-kit  
 RT (mast/stem cell growth factor receptor) proto-oncogene in human  
 RT piebaldism.";  
 RL Am J. Hum. Genet. 50:261-269(1992).  
 RN  
 [6] VARIANT ARG-664.  
 CC MEDLINE-92020918; PubMed-1717985;  
 RA Giebel L.B., Spritz R.A.;  
 RT "Mutation of the KIT (mast/stem cell growth factor receptor)  
 RT proto-oncogene in human piebaldism.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:8696-8699(1991).  
 RN  
 [7] VARIANT VAL-816.  
 CC MEDLINE-94013473; PubMed-7691885;  
 RA Furitsu T., Tsujimura T., Tono T., Ikeda H., Kitayama H.,  
 RA Koshinzu U., Sugahara H., Butterfield J.H., Ashman L.K.,  
 RA Kanayama Y., Matsuzawa Y., Kitamura Y., Kanakura Y.;  
 RT "Identification of mutations in the coding sequence of the proto-  
 RT oncogene c-kit in a human mast cell leukemia cell line causing  
 RT ligand-independent activation of c-kit product.";  
 RL J. Clin. Invest. 92:1736-1744(1993).  
 RN  
 [8] VARIANTS PIEBALDISM GLY-791 AND VAL-812.  
 CC MEDLINE-93322624; PubMed-7687267;  
 RA Spritz R.A., Holmes S.A., Itin P., Kuester W.;  
 RT "Novel mutations of the KIT (mast/stem cell growth factor receptor)  
 RT proto-oncogene in human piebaldism.";  
 RL J. Invest. Dermatol. 101:22-25(1993).  
 RN  
 [9] VARIANT PIEBALDISM 893-GLU--PRO-896 DEL.  
 CC MEDLINE-96287384; PubMed-8680409;  
 RA Riva P., Milani N., Gandolfi P., Larizza L.;  
 RT "A 12-bp deletion (7818del12) in the c-kit protooncogene in a large  
 RT Italian kindred with piebaldism.";  
 RL Hum. Mutat. 6:343-345(1995).  
 RN  
 [10] VARIANT GIST VAL-559 DEL.  
 CC MEDLINE-98361155; PubMed-9697690;  
 RA Nishida T., Hirota S., Taniguchi M., Hashimoto K., Isozaki K.,  
 RA Nakamura H., Kanakura Y., Tanaka T., Takabayashi A., Matsuda H.,  
 RA Kitamura Y.;  
 RT "Familial gastrointestinal stromal tumours with germline mutation of  
 RT the KIT gene.";  
 RL Nat. Genet. 19:323-324(1998).  
 CC  
 CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL  
 CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.  
 CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND  
 CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL  
 CC 3-KINASE (PI3K).  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- DISEASE: DEFECTS IN KIT ARE THE CAUSE OF A DEVELOPMENTAL  
 CC ABNORMALITY KNOWN AS PIEBALDISM. AN AUTOSOMAL DOMINANT GENETIC  
 CC DISORDER OF PIGMENTATION CHARACTERIZED BY CONGENITAL PATCHES OF  
 CC WHITE SKIN AND HAIR THAT LACK MELANOCYTES.  
 CC -!- DISEASE: DEFECTS IN KIT ARE THE CAUSE OF A GASTROINTESTINAL  
 CC STROMAL TUMOR (GIST).  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD117 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd117.htm";  
 CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;  
 CC WWW="http://www.infobio.gen.fr/services/chromancer/genes/KITID127.html".  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC DR EMBL; X06182; CAA29548.1; -;  
 CC DR EMBL; X69301; CAA49159.1; -;  
 CC DR EMBL; X69302; CAA49159.1; JOINED.  
 CC DR EMBL; X69303; CAA49159.1; JOINED.  
 CC DR EMBL; X69304; CAA49159.1; JOINED.  
 CC DR EMBL; X69305; CAA49159.1; JOINED.  
 CC DR EMBL; X69306; CAA49159.1; JOINED.  
 CC DR EMBL; X69307; CAA49159.1; JOINED.  
 CC DR EMBL; X69308; CAA49159.1; JOINED.  
 CC DR EMBL; X69309; CAA49159.1; JOINED.  
 CC DR EMBL; X69310; CAA49159.1; JOINED.  
 CC DR EMBL; X69311; CAA49159.1; JOINED.  
 CC DR EMBL; X69312; CAA49159.1; JOINED.  
 CC DR EMBL; X69313; CAA49159.1; JOINED.  
 CC DR EMBL; X69314; CAA49159.1; JOINED.  
 CC DR EMBL; X69315; CAA49159.1; JOINED.  
 CC DR EMBL; X69316; CAA49159.1; JOINED.  
 CC DR EMBL; X69317; CAA49159.1; JOINED.  
 CC DR PIR; S01426; TVHUKT.  
 CC DR HSSP; P11362; IFGK.  
 CC DR Genew; HGNC:6342; KIT.  
 CC DR MIM; 164920; -;  
 CC DR MIM; 172800; -;  
 CC DR MIM; 606764; -;  
 CC DR InterPro; IPR000719; Euk\_pkinase.  
 CC DR InterPro; IPR003006; Iq\_MHC.  
 CC DR InterPro; IPR003598; Iq\_c2.  
 CC DR InterPro; IPR001824; RTKinaseII.  
 CC DR InterPro; IPR001245; Tyr\_pkinase.  
 CC DR Pfam; PF00047; Iq; 1.  
 CC DR Pfam; PF00069; pkinase; 1.  
 CC DR ProDom; PD000001; Euk\_pkinase; 2.  
 CC DR SMART; SM00408; IGC2; 1.  
 CC DR SMART; SM00219; TyKc; 1.  
 CC DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 CC DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 CC KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;  
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;  
 KW Immunoglobulin domain; Disease mutation.  
 FT SIGNAL 1 22  
 FT CHAIN 23 976 MAST/STEM CELL GROWTH FACTOR RECEPTOR.  
 FT DOMAIN 23 520 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 521 543 POTENTIAL.  
 FT DOMAIN 544 976 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 589 937 PROTEIN KINASE.  
 FT NP\_BIND 595 603 ATP (BY SIMILARITY).  
 FT BINDING 623 623 ATP (BY SIMILARITY).  
 FT ACT\_SITE 792 792 BY SIMILARITY.  
 FT MOD\_RES 823 823 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 130 130 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 145 145 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 283 283 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 293 293 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 300 300 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 320 320 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 352 352 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 367 367 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 463 463 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 486 486 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT VARIANT 559 559 MISSING (IN GIST).  
 FT VARIANT 583 583 /FTID=VAR\_007965.  
 FT VARIANT 583 583 E -> K (IN PIEBALDISM).  
 FT VARIANT 584 584 F -> L (IN PIEBALDISM).  
 FT VARIANT 664 664 /FTID=VAR\_004105.  
 FT VARIANT 664 664 G -> R (IN PIEBALDISM).



```
FT FT 791 791 /FTID-VAR_004106.
FT R -> G (IN PIEBALDISM).
FT FT 812 812 /FTID-VAR_004107.
FT G -> V (IN PIEBALDISM).
FT FT 816 816 /FTID-VAR_004108.
FT D -> V (IN MAST CELL LEUKEMIA;
FT CONSTITUTIVELY ACTIVATED).
FT FT 893 896 /FTID-VAR_004109.
FT MISSING (IN SEVERE PIEBALDISM).
FT FT 976 976 /FTID-VAR_004110.
FT /FTID-VAR_004110.
SQ SEQUENCE 976 AA; 109864 MW; 81B0CD76817F3454 CRC64;

Query Match 23.7%; Score 1251; DB 1; Length 976;
Best Local Similarity 32.7%; Pred. No. 3.8e-76;
Matches 324; Conservative 166; Mismatches 352; Indels 148; Gaps 34;

QY 47 VKSSSYPMWSESPDLGCLALRQSSGTVYEAARVVDVNSASTLQVLVDAPGNISCLWV 106
DB 20 VQTGSSQPSVSPG-EPSPSPHFGKSDLI-----VRVGEIRLLCTDPCFVK--WT 67
QY 107 FKHSSLCOPHFDLQNRGVVSMVLAKWTQAGEYLLFIOSEATN---YTLLETVSIRNT 163
DB 68 FE-----ILDETENKQNEWIT-EKAEATNTCKYTCTNNKHGLSNS 106
QY 164 L-----LYTLRPYFRKMENQDALVCSISVPEPIVWVWLCDSQSGESCKEE----- 209
DB 107 IYVFPDPAKFLVDRSLYCK-EDNTLVRCPLTDPE-VTNYSLKCGQKPLPKDLRFIP 164
QY 210 ----SPAVKKEKVLHELFETDIRCCARNELGRECTRLFTIDL-----NOTPQTLPQ--L 260
DB 165 DPKAGIMIKSVKRAYHRLC---LHCSVDQCKSVLSEKFLTKVRPAKPAVVPVSVSKASY 221
QY 261 FLKVGCEPLIRKAVHNGFGLTWLEN---KALBEGNYFENSTVSTNTRMIRILFAFV 317
DB 222 LREGEETFTCTIKDVSSVSTWRENSQTKLOEKYNSWHHGDGFNYEQQAT-----LTI 277
QY 318 SSVARNDTGYTCCSSKHPQSALVT---IVGKGFNA-----TNSSEYDEIDQYE 365
DB 278 SSARVNDSGVFCMYANTGCSANVTITLWVDKGFNFIPMINTVTFVNDGENVDL----- 333
QY 366 EFCFSVFRKAYPO-ICTWTFSTRKSF--PCEOKGLONGYSISKFNHKH-----OPGE 415
DB 334 ----IVEYEAFFPKPEHQWIYMNRTDWDYPKSENEISNIRYVSELHLTRLKGTBGGT 389
QY 416 YIFHAENDAAQTKMTLAIIRKPOVLA--EASASQASCFSDGYPPLSPWTKWCKSDKSPN 473
DB 390 YTLVSNDSVNAIAFNVTYNTKPEILTYDRLVNGMLQCVAAAGFPEPTIDWYFCPCTEQR 449
QY 474 CTEETE-GVWNRKANRKYFGOWSVSSSTLNMSEAIGFLKCCAYNSLGTSCETILLNSP 532
DB 450 CSASVLPVDVQTLNNSGPPFGKLVQSSIDSSAFKNGIVECKAYNDVGT--SAYFN-- 505
QY 533 GPFPFTQDN-----ISFYATIGVCLLFTVVLTLILCHIKYKQFRIESQLQM 578
DB 506 --FAFGKNNKEQTHPTLFTPLLIGFVIVAGMCIIVMLT-----YKLOKPMYEQWKV 559
QY 579 V-QVTGSSDNEFYVDREYEDLKEFPRENLEPKVILGSGAFGKVMATVIGSKTV 637
DB 560 VEEING---NNVYIDPTQLPYDHKEFPNRLSFGKTLGAGAFGVWEATAYGLIKSDA 616
QY 638 STQAVKMLKEKADSRERALSSELKMMPTOLGSHENIVNLGACTLSGYLLIFYECCYV 697
DB 617 AMTVAVKMLKPSAHLTERALMSELKLVSLGNHNMIVNLGACTIGGPTLVITYECCYV 676
QY 698 DLLNLYLRKREKHEFTWTE-----IFKEHNFSEYPTFSQHPNSSM---PGSREVOIHPD 748
DB 677 DLLNLYLRKREKHEFTWTE-----IFKEHNFSEYPTFSQHPNSSM---PGSREVOIHPD 748
QY 749 SDQISGLHNSHSEDEIYENOKRLEBEDENLVTFEDLLCFAYQAVAKMGFEKFSKV 808
DB 736 ADKRRSVRIGSY-----IERDVTYPAIMEDDEL-ALDLELLSFSYQAVKMAFLASKNCI 789
QY 809 HRDLARNVLVTHGKVVVKICDFGLARDINSDSNVYVGRNARLPVKWMAPESLFEGYITIK 868
```

```
DB 790 HRDLARNVLLTHGRITKICDFGLARDIKDSDNYVVKGNARLPVKWMAPESEIFNCVTFE 849
QY 869 SDVWSYGILLWEIFSLGVNPPGIPVDANFYKLTONGKMDQPPFYATEEYIIMQSCHAF 928
DB 850 SDVWSYGIFLWEIFSLGSSPYGMPVDKSKFYKMKIEGFRMLSPHAPAEYDIMKTWDA 909
QY 929 DSRKRPSFPNLTSLGCOLADAEEMAYQNV 958
DB 910 DPLKRPTEFKQIVQLIEKQISESTNHIYSNL 939

RESULT 7
KIT_CANFA
ID KIT_CANFA STANDARD; PRT; 975 AA.
AC O97799;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).
GN KIT.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99142897; PubMed=9989791;
RA Ma Y., Longley B.J., Wang X., Blount J.L., Langley K., Caughey G.H.;
RT "Clustering of activating mutations in c-KIT's juxtamembrane coding
region in canine mast cell neoplasms.";
RL J. Invest. Dermatol. 112:165-170(1999).
CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL: AF044249; AAD02327.1;
DB HSSP: P11362; 1FGK.
DB InterPro: IPR000719; Euk_pkinase.
DB InterPro: IPR003599; Ig.
DB InterPro: IPR003006; Ig_MHC.
DB InterPro: IPR003600; Ig_Like.
DB InterPro: IPR001824; RTKinaseIII.
DB InterPro: IPR001245; Tyr_pkinase.
DB Pfam: PF00047; Ig; 2.
DB Pfam: PF00069; pkinase; 1.
DB ProDom: PD000001; Euk_pkinase; 2.
DB SMART: SM00409; IG; 2.
DB SMART: SM00410; IG_Like; 1.
DB SMART: SM00219; TyrcK; 1.
DB PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DB PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DB PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
DB PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
```

Immunoglobulin domain.  
KW SIGNAL 1 24  
FT CHAIN 25 975  
FT DOMAIN 25 975  
FT TRANSMEM 520 542  
FT DOMAIN 543 975  
FT DOMAIN 588 936  
FT NP\_BIND 594 602  
FT BINDING 622 622  
FT ACT\_SITE 791 791  
FT MOD\_RES 822 822  
FT CARBOHYD 96 96  
FT CARBOHYD 132 132  
FT CARBOHYD 147 147  
FT CARBOHYD 286 286  
FT CARBOHYD 296 296  
FT CARBOHYD 303 303  
FT CARBOHYD 355 355  
FT CARBOHYD 370 370  
FT CARBOHYD 403 403  
FT CARBOHYD 466 466  
FT CARBOHYD 489 489  
FT CARBOHYD 489 489  
SQ SEQUENCE 975 AA; 109335 MW; 8F570BDB9F05B1CB CRC64;

Query Match 23.7%; Score 1247.5; DB 1; Length 975;  
Best Local Similarity 33.7%; Pred. No. 6.5e-76;  
Matches 333; Conservative 161; Mismatches 346; Indels 149; Gaps 39;

QY 47 VKKSSPYMVSPEDELCALRQSSGTVEAAVEVDVSASITLQVLVDAPGNISCLWV 106  
DB 22 VRTGSSQPSVSGPSLIP-SIHPAKS-----ELIVSGDELRLSCDTPGFKV-WT 69  
QY 107 FKX-SSLNCPHEDLQNRGVSVILKMTQAGEYLLFIQSSEATNYTIL-----FT 157  
DB 70 FETLQGNLNTHT-----NEWITE-----KAEAG-----HTGNYTCTNRDGLRSIY 110  
QY 158 VSTRN-TLTYLRPRYFRKMNEDALVISEVPEPIVEWVLCDSDGCKEESPAV--- 213  
DB 111 VVRDPKALFLDLPLYGK-EGNDTLVRCLTPDE-VTNSLRGCEGKPLPKDLTFVADP 168  
QY 214-----VKKEKVLHFGTDIRC--CARNELGRE-CTRLETFID-----LNQTPQTLPLQ- 260  
DB 169 KAGITIRNVKREHRL-----CLHCSADQKGRVLSKFTLVKRAIRAVPVVSVSKTS 222  
QY 261 -FLKVGCEPLWIRKAVVNHGFLTWELEN-----KALEEGNYFEMSTYSTNTRWIRLFA 315  
DB 223 SLLKEGAFSVMCFIKDVSVFSVDMWIKENSQQTNAQTQSNWHHGDFFNERQEKLI--- 279  
QY 316 FVSSVARNTGYTCSSSKHPSQSALVT---IVGKGFINA-----TNSSEYEDIQ 363  
DB 280 -ISSARVNSGVPCVANNFTGSGANVTTLLEVVDKGFINIFPMNMSTIFVNDGQVNDL-- 336  
QY 364 YEEFCFSVRKAYPO-IRCTWTFSRKSFQCEQKGL---DNQYSISKFCNHHK-----Q 412  
DB 337 -----IVEEAYPKPEHQWIWMNFTFTDKWEDYPKSDNESNI-RYVSELHLTRLKGN 389  
QY 413 PGYIFHAENDDAQFTKFTLNRKPOVLAEASQAA--SCFSDGYPLPSWTWKCSDK 470  
DB 390 GCTYTFQVNSDYNSSVTFVNYVNTKPEILTHESLNGMLQCVVAGFPFAVGYFCPGA 449  
QY 471 SPNCTEEI-TEGVNKKANKKVFQGWVSSSTLNNSEAIKGLFKVCCAYNSLGTSCETILL 529  
DB 450 EQRCSVPIDPMYQMNQSSLSPSGKLVQSSIDYSAFKHNGTVECRAYNVNVRGS--SAFF 507  
QY 530 NSPQPPFFIQD-----NISFYATIGVCLLFIIVLLTLLCHYKKKFRYESQLQMV 579  
DB 508 N-----FAFKEQIHPHTLFTPLLIGFVIAAGMCMCIIVMLT-----YKYLQKPMVEYQWVV 559  
QY 580 -QVTGSDNDEYFVDRPREYEDYDLKWEPPRENLEFGKVLGSGATKVMNATAYGISKTGVS 638  
DB 560 EEING---NNYVYIDTQLPYDHYKHKEFPNRLRSLFGKTLGAGAFCKVVEATAYGLIKSDAA 616  
QY 639 IQVAVKMLKEKADSSEREALMSSELKMMTQLGSHENIVNLLGACTLSGPIYLIFEYCCYGD 698

DB 617 MTVAVKMLKPSAHLTEREALMSSELKLVSLYLGNNHMINVLLGACTVGGPTLVITEYCCYGD 676  
QY 699 LNYLSKREKFRHTWTE-----IFKEHNFSYPTFQSHPNSSM---PGSREVQIHPDS 749  
DB 677 LNFLRRKRDSFCSKQEDHGEVALYKNNLLHSSKSSDSTNEYMDMKPGVSIV-VPTKA 735  
QY 750 DOIISGLHNSHSEDETEYENQKRLSEEDLNVLTPEDLCLCFAYOVAKGMEFLFKSCVH 809  
DB 736 DKRSRARIQS-----IERDVTFAIMEDDEL-ALDLEDLLSFSYQVAKGMAFLASKNCIH 789  
QY 810 RLDAARNLVTHGKVVVKVVICDFGLARDIMSDSNVYVRGNARLPVKWMAPELSLFEQIYTIKS 869  
DB 790 RLDAARNILLTHGRITKICDFGLARDIKDNDNSYVYKGNARLPVKWMAPEISFNCVYTFES 849  
QY 870 DVWSYGIILLWEISLGVNYPGIPVDANFYKLTQNGFKMDQPPFYATEEIIYIMOSCAFD 929  
DB 850 DVWSYGIIFLWELFSLGSSPYGPMVDKSKFYKMIKEGFMLSPEHAPAEYMDIMKTCWAD 909  
QY 930 SRKRPSFPNLTSLFGQLADAEAEAMQNV 958  
DB 910 PLARPTFKQIVQLIEKQISDSTNHIYSNL 938

RESULT 8  
KEMS\_FELCA STANDARD; PRT; 980 AA.  
AC P13369;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)  
DE (EC 2.7.1.12) (fms proto-oncogene) (c-fms).  
GN CSF1R OR FMS.  
OS Felis silvestris catus (Cat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
OX NCBI\_TaxID:9685;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-89077553; PubMed-2849512;  
RA Woolford J., McAuliffe A., Rohrschneider L.R.;  
RT "Activation of the feline c-fms proto-oncogene: multiple alterations  
are required to generate a fully transformed phenotype.";  
RL Cell 55:965-977(1988).  
CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN  
TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.  
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
tyrosine phosphate.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
PROTEIN KINASES.  
CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: J03149; AAA30811.1; -;  
DR PIR: A31636; TVCTMD.  
DR HSSP: P11362; 1FGK.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR003598; Ig\_C2.  
DR InterPro: IPR003600; Ig\_Like.  
DR InterPro: IPR001824; RtkKinaseIII.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR Pfam: PF00047; Ig; 3.  
DR Pfam: PF00069; pkinase; 1.

DR ProDom; PD000001; Euk\_pkinase; 2.  
 DR SMART; SM00410; IG\_like; 4.  
 DR SMART; SM00408; IGc2; 1.  
 DR SMART; SM00219; TyrKc; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;  
 KW Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;  
 KW Immunoglobulin domain; Repeat.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 980 MACROPHAGE COLONY STIMULATING FACTOR I  
 FT RECEPTOR.  
 FT DOMAIN 20 509 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 510 535 POTENTIAL.  
 FT DOMAIN 536 980 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 24 104 IG-LIKE C2-TYPE DOMAIN 1.  
 FT DOMAIN 107 197 IG-LIKE C2-TYPE DOMAIN 2.  
 FT DOMAIN 204 298 IG-LIKE C2-TYPE DOMAIN 3.  
 FT DOMAIN 299 397 IG-LIKE C2-TYPE DOMAIN 4.  
 FT DOMAIN 398 502 IG-LIKE C2-TYPE DOMAIN 5.  
 FT DOMAIN 579 908 PROTEIN KINASE.  
 FT NP\_BIND 585 593 ATP (BY SIMILARITY).  
 FT BINDING 613 613 ATP (BY SIMILARITY).  
 FT ACT\_SITE 776 776 BY SIMILARITY.  
 FT DISULFID 42 84 POTENTIAL.  
 FT DISULFID 127 177 POTENTIAL.  
 FT DISULFID 224 278 POTENTIAL.  
 FT DISULFID 417 482 POTENTIAL.  
 FT MOD\_RES 807 807 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 45 45 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 73 73 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 94 94 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 153 153 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 275 275 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 286 286 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 302 302 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 335 335 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 410 410 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 477 477 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 490 490 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 980 AA; 108506 MW; 4E5CF61E97CFCEFF CRC64;  
 Query Match 23.2%; Score 1225.5; DB 1; Length 980;  
 Best Local Similarity 33.5%; Pred. No. 1.9e-74;  
 Matches 321; Conservative 168; Mismatches 321; Indels 149; Gaps 34;  
 QY 83 VDVSASITQLVLVDAPGNISCLW---VFKHSLNCP-----HFDLQNRGVVSMVIL 131  
 DB 32 VEPGTVTLRCV---GNGSVEDWGPISPHWNLDLDPSSILTTNNATFONGTYHCT-- 85  
 QY 132 KMTQAGEYL--LFIQSEATNYTL---FTYSIRNTLLTYLRPFYFRKMENQDALV-CI 185  
 DB 86 EPGNPGGNATHLVKVPDPAPKPKVLAQEVTV-----LEGODALLPCL 128  
 QY 186 SESVPEPVE--WVLCDQSGESCKEE-----SP---AVVKEEKVLHFLPGTDIRCCARN 235  
 DB 129 ---LDPDALEAGSVLRVGRVRLQNTYFSFPHGFTTHKAKFTENHVY---QCSARV 181  
 QY 236 ELGRCETRL-----FTIDLNPTQTLT-PQLFLKV-GEPLWIRCKAVHVNHGFGTLWELE 288  
 DB 182 D-GRVTWNGWLKVKQDKISGPATLLEPAELVRTOGEAAQVCSASINVDVDFV----- 235  
 QY 289 NKALEGN-----YFEMSTYSTNRTMIRLFAFVSSVARNDTGYTCSSSK---HPSQSAL 341  
 DB 236 --SLRHGDTKLTISSQSDFDHNRQY-KVLTNLHDVHVSFODAGNYSCTAINANGNSASVY 292  
 QY 342 VTIVGKGGINATN--SSEYDEIDQEEFCFSVRFKAYPQIRCT--TWTFSRKSFCEQKGLD- 398  
 DB 293 FRVVESAYLNTLSEQLLQEVTVGKVDLQVKVEAYPGLESPFNWTV-LGPFSDYQDKLDF 351  
 QY 399 ---NGYSISKFCN---HKHPGGEYFHAENDDAQFTKMTFLNTRKPKQVLAESASQA 450

DB 352 VTIKDYRYVTSTLSLPRKLRSEAGRYSFSLARNAGGQWALTFELTLRYPPVRYVTMTLNG 411  
 QY 451 S-----CFSDGYPLPSTWTKKCDKSPNCTEE---ITEGVNKRANKRVKFCQWSSSTLNM 503  
 DB 412 SDTLCLCEASGYPOPSVTWQCRSHTRDCDESAGLVLEDSEVLSQSPFPFHEVIVHSLAI 471  
 QY 504 SEAIKGFELVKCCAYNSLGSCTETILLNSPGPPFFIDONISFYATIGVC---LLFIVVLT 560  
 DB 472 GTLEHNRTECFRAFNSVGNSSQTFWFIPIGAHTQLDPDELLFTPVLLTCMIMALLULLL 531  
 QY 561 LICHYKKKOFRYESQLQMVQVQTGSSDNEYFYDFREYEDLKWEFFRENLEFGKVLGSGA 620  
 DB 532 LLLYKQKPKYQVRWKIIE--SYEGNSYTFIDPTOLPYNEKWEFFPENNLOFGKTLGAGA 589  
 QY 621 FGKVMNATAGISKTGVSTQAVKMLKEKADSSERALMSELKMMTOLGSHENIVNLLGA 680  
 DB 590 FGKVEATAFGLKEDAVLKAVKMLKSTAHADKEKALMSELKIMSHLGQHENIVNLLGA 649  
 QY 681 CTLSGPIYLIFEYCCYCDLLNLYLSKRE-----KEHRTWTEIFKE 720  
 DB 650 CTGGGPVLVITEYCCYCDLLNLRQAEAMLGPSLSVGQDPEAGAGYKNHLEKKYVRD 709  
 QY 721 HNFSEYPTFQSHPNSSMPGSRVQIHPDSQISGLHGNSEHSEDEIEYENOKRLEEEEDL 780  
 DB 710 SDPS-----SOGVDYTVEMRPVSTSSNDSFSEEDLCKEDGRPLE----- 749  
 QY 781 NVLTFEDLLCFAYQVAKGMEFFLEFKSCVHRDLAARNVLTHTGKVVKTCDFGLARDIMSDS 840  
 DB 750 ---LRDLLHFSQVAGQMAFLASKNCIHRDVAARNVLTSGRVAKIGDFGLARDIMNDS 805  
 QY 841 NYVVRNARLPVKWMAPESEFEGIIYIKSDVWSYGILLWEIFSLGVNPNYPGIPVDANFYK 900  
 DB 806 NYIVKNARLPVKWMAPESEFDCVYTVQSDVWSYGILLWEIFSLGLNYPGILVNSFYK 865  
 QY 901 LIONGFKMDPPFYATEIIVIMOSWAFDSKRPSPNLTSLGCOL-ADAEAMYONV 958  
 DB 866 LVKDGQVMAQPAFAFPKNIYSIMOACWALEPRTTQQICSLQKQAEQDRRVPNYTNL 924  
 RESULT 9  
 KIT\_FELCA STANDARD; PRT; 978 AA.  
 AC Q28889;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)  
 DE (Proto-oncogene tyrosine-protein kinase kit) (c-kit).  
 GN KIT.  
 OS Felis silvestris catus (Cat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
 OX NCBI\_TaxID=9685;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Head;  
 RX MEDLINE=95140426; PubMed=7530827;  
 RA Herbst R., Munemitsu S., Ullrich A.;  
 RT "Oncogenic activation of v-kit involves deletion of a putative  
 RT tyrosine-substrate interaction site.";  
 RL Oncogene 10:369-379(1995).  
 CC -!- FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL  
 CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.  
 CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND  
 CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL  
 CC 3-KINASE (PI3K).  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; S76596; AAB33207.1; -  
 CC HSSP; P11362; IFKC.  
 CC InterPro; IPR000719; Euk\_pkinase.  
 CC InterPro; IPR003599; Ig.  
 CC InterPro; IPR003006; Ig\_MHC.  
 CC InterPro; IPR003600; Ig-like.  
 CC InterPro; IPR001824; TykinaseIII.  
 CC InterPro; IPR001245; Tyr\_pkinase.  
 CC Pfam; PF00047; Ig; 1.  
 CC Pfam; PF00069; pkinase; 1.  
 CC ProDom; PD000001; Euk\_pkinase; 2.  
 CC SMART; SM00409; IG; 2.  
 CC SMART; SM00410; IG\_like; 1.  
 CC SMART; SM00219; TyrKc; 1.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 CC PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 CC ProTo-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;  
 CC Transferrase; Glycoprotein; Phosphorylation; ATP-binding; Signal;  
 CC Immunoglobulin domain.  
 CC SIGNAL 1 22  
 CC CHAIN 23 978  
 CC DOMAIN 23 521  
 CC TRANSFEM 522 544  
 CC DOMAIN 545 978  
 CC DOMAIN 590 938  
 CC NP\_BIND 596 604  
 CC BINDING 624 624  
 CC ACT\_SITE 793 793  
 CC MOD\_RES 824 824  
 CC CARBOHYD 94 94  
 CC CARBOHYD 145 145  
 CC CARBOHYD 284 284  
 CC CARBOHYD 294 294  
 CC CARBOHYD 301 301  
 CC CARBOHYD 321 321  
 CC CARBOHYD 353 353  
 CC CARBOHYD 368 368  
 CC CARBOHYD 401 401  
 CC CARBOHYD 464 464  
 CC CARBOHYD 487 487  
 CC SEQUENCE 978 AA; 109449 MW; 6D45472E07440E6B CRC64;

Query Match 23.2%; Score 1224.5; DB 1; Length 978;  
 Best Local Similarity 32.9%; Pred No. 2.3e-74;  
 Matches 325; Conservative 167; Mismatches 352; Indels 145; Gaps 35;  
 QY 47 VGGSSYPMVSEPDGCG-ALRPOSSGTYEAAAVEVDVSAITLQVLVADPAGNISCLW 105  
 DB 20 VQTGSSQP--SASPGWSLPSIHPTSPS-----ELIVSAGDEIRLLCTDPGFVK--W 66  
 QY 106 VFKHSLNCPHFQDLQNRGVSMVILKWTETQAGEVLLFIQSEAT---NYTIL----- 155  
 DB 67 TFE-----TLGQSSEITHNEWIT-EKAEATNTGNYCTNGGGLSS 105  
 QY 156 -FTVSTIRN-TLLYTLRRPYFRKMNODALVCISESPPEIVENVLCDSDGCKEESPAV 213  
 DB 106 STYVFRDPAKFLVDLPYLGK-EDHDTLVRCPLTDPE-VTVNSLRGCGKPLPKDLTFV 163  
 QY 214 -----VKKEKVLHFGTDIRCCA-----RNLGRECTRLFTIDLQTPOTTLPOL- 260  
 DB 164 TDPKAGITIRNVKREYHRLC---LHCSADRKCKSVLSKFKTLKVRRAIRAVPVVSYSKAS 220

QY 261 -FLKVGEPWIRCKAVHVNHGFGTLWELENKALEGNYFEMSTYSTNRTMIRILFAFVS 319  
 DB 221 HLLREGEESVWCLIKDVSSSVSDSMWIKENSPOTNAQPSQNSHQGDENFVROERUTISS 280  
 QY 320 VARNDTGYTCSSSHPSOSALVT---IVKGKGPINA-----TNSEDDYIDQIEEF 367  
 DB 281 ARVNSDGVFMCYANNTFGSANVTTTLEVAKGFIPFMMNTTIFVNDGNGVDL----- 334  
 QY 368 CFSVRFKAYPQ-----IRCTWTFSRKSFCEQKGLDNGYSISKFCNKH-----Q 412  
 DB 335 --IVEYEAYPKPEHQWVYNNRSLTLDKWDYPKS-----DNESNI-RVYSELHLTRLKGN 387  
 QY 413 PGYEIFHAENDDAQFTKMETLNIRRRPQVLAELAS--ASQACSFSDGYPLPSWTWKKCSDK 470  
 DB 388 GGYTFQVSNDSVNSVTLNVYVNTKPEILTHESVSGILQCLVAGPEPFDVWYFCPGA 447  
 QY 471 SPNCTEITE-GVMNRKANRKFQWVSSSTLNMSRAIKGFLYKCCAYNSLGTSCETILL 529  
 DB 448 EQRCPPVPGLDVQMNSVSPSGKLVSQSIDYSAFKHNGTVECRASNNVGKT--SAFF 505  
 QY 530 NSPGPPFFTIQDN-----ISFYATIGVCLLFIWVLTLLICHKKYKQFRYESQ 575  
 DB 506 N----FAFKGNSKEQMHPTLFTPLLJIGFVIAAGMCIIVMILT---YKLOKPMTEVQ 557  
 QY 576 LQMV-OVTGSSDNEYFYVDREYEDLKWEFPRENLEFGLGVLGSGAPKVMNATAYGISK 634  
 DB 558 WKVVEEING---NNYVYIDTQPLPYDHKWEFPNRLSFGKTLGAGAFKVVETATAYGLIK 614  
 QY 635 TGVSIQVAVKMLKEKADSSEREALMSKMTQLGSHENIVNLLGACTLSGPIYLIFEYC 694  
 DB 615 SDAAMTVAVKMLKPSAHLTEREALMSKLVLSYLNHNMIVNLLGACTVGGPTLVITEYC 674  
 QY 695 CYGDLNLYLRKREKFRHTWTEFKENHSFYPT-FQSHPNSSMPGSGRE-VQIHPDSQI 752  
 DB 675 CYGDLNLFRRKRDSFICSQEDHAE--VALYKNLLQSKESSCNDSTNEYMDMKPGVSYV 732  
 QY 753 SGLHGSFSE---DETEYENQKRLEEDDLNLTREDLLCFAYQVAKGMEFLFEKSCVH 809  
 DB 733 VPTKADRRKSARIGSYIERDVTPTAIMEDEL-ALDLELLSFQVAKGMAFLASKNCIH 791  
 QY 810 RLDAARNLVTHGKVKICDFGLARDIMSDSNVYVGNARLPVKWMAPELSFEGIYTIKS 869  
 DB 792 RLDAARNLLTHGRITKICDFGLARDIKDSNVYVGNARLPVKWMAPELSFNCVYTFES 851  
 QY 870 DWYSYGILLWEISLGVNYPGIPVDANFYKLQNGFKMDQPPYATETIYLIMOSCHAFD 929  
 DB 852 DWYSYGIFLWEISLGVNYPGIPVDANFYKLQNGFKMKIEGFRMLSPHAPAEYMDIMKTCWDAD 911  
 QY 930 SRKRPSFPLNTSLFGCOLADAEAEAMQNV 958  
 DB 912 PLKRPTFKQIVQLIEKQISDSTNHIYSNL 940  
 RESULT 10  
 KMS\_FSVMD  
 ID KMS\_FSVMD STANDARD; PRT; 978 AA.  
 AC P00545; Q86597;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tyrosine-protein kinase transforming protein fms (EC 2.7.1.112).  
 GN V-FMS.  
 OS Feline sarcoma virus (strain McDonough).  
 OC Viruses; Retroid viruses; Retroviridae; Mammalian type C retroviruses.  
 OX NCBI\_Taxid=11778;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84119469; PubMed=6582485;  
 RA Hampe A., Gobet M., Sherr C.J., Galibert F.;  
 RT "Nucleotide sequence of the feline retroviral oncogene v-fms shows  
 RT unexpected homology with oncogenes encoding tyrosine-specific protein  
 RT kinases.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 81:85-89(1984).

[2]  
 REVISIONS, SEQUENCE FROM N.A.  
 MEDLINE=92015516; PubMed=1833563;  
 RA Smola U., Hennig D., Hadwiger-Fangmeier A., Schuetz B., Pfaff E.,  
 RA Niemann H., Tanura T.,  
 RT "Reassessment of the v-fms sequence: threonine phosphorylation of the  
 COOH-terminal domain";  
 RL J. Virol. 65:6181-6187(1991).  
 CC -!- FUNCTION: V-FMS IS DERIVED FROM THE RECEPTOR FOR COLONY  
 CC STIMULATING FACTOR 1 (CSF-1).  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-FMS  
 CC POLYPEPTIDE.  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: K01643; AAA43045.1;  
 CC EMBL: S59588; AAB20028.1;  
 CC PIR: A00654; TMVMD.  
 CC HSP: P11362; IFGK.  
 CC InterPro: IPR000719; Euk\_pkinase.  
 CC InterPro: IPR003006; Ig\_MHC.  
 CC InterPro: IPR003598; Ig\_C2.  
 CC InterPro: IPR003600; Ig\_Like.  
 CC InterPro: IPR001824; RtkinaseII.  
 CC InterPro: IPR001245; Tyr\_pkinase.  
 CC Pfam: PF00047; Ig; 3.  
 CC Pfam: PF00069; pkinase; 1.  
 CC ProDom: PD000001; Euk\_pkinase; 2.  
 CC SMART: SM00410; Ig\_Like; 4.  
 CC SMART: SM00408; IGC2; 1.  
 CC SMART: SM00219; TYRC; 1.  
 CC PROSITE: PS00107; PROTEIN\_KINASE\_APP; 1.  
 CC PROSITE: PS00109; PROTEIN\_KINASE\_TVR; 1.  
 CC PROSITE: PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 CC PROSITE: PS0011; PROTEIN\_KINASE\_DOM; 1.  
 KW Polyprotein; Tyrosine-protein kinase; Oncogene; Transferase; Receptor;  
 KW Transmembrane; Glycoprotein; Phosphorylation; ATP-binding;  
 KW Immunoglobulin domain; Repeat.  
 FT DOMAIN 1 543 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 544 POTENTIAL.  
 FT DOMAIN 569 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 58 138  
 FT DOMAIN 141 231 IG-LIKE C2-TYPE DOMAIN 1.  
 FT DOMAIN 238 332 IG-LIKE C2-TYPE DOMAIN 2.  
 FT DOMAIN 333 431 IG-LIKE C2-TYPE DOMAIN 3.  
 FT DOMAIN 432 536 IG-LIKE C2-TYPE DOMAIN 4.  
 FT DOMAIN 613 942 IG-LIKE C2-TYPE DOMAIN 5.  
 FT NP\_BIND 619 PROTEIN KINASE.  
 FT BINDING 647 ATP (BY SIMILARITY).  
 FT DISULFID 76 118 POTENTIAL.  
 FT DISULFID 161 211 POTENTIAL.  
 FT DISULFID 258 312 POTENTIAL.  
 FT DISULFID 451 516 POTENTIAL.  
 FT MOD\_RES 841 973 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 79 79 PHOSPHORYLATION.  
 FT CARBOHYD 107 107 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 128 128 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 187 187 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 309 309 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 320 320 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 336 336 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 369 369 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 444 444 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 511 511 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 524 524 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CONFLICT 714 714 L -> P (IN REF. 1).  
 FT CONFLICT 971 978 QRTPPVAR -> RGPPL (IN REF. 1).  
 SQ SEQUENCE 978 AA; 108491 MW; 4C7CAC4835185EBF CRC64;  
 Query Match 23.28; Score 1224; DB 1; Length 978;  
 Best Local Similarity 32.6%; Pred. No. 2.4e-74;  
 Matches 334; Conservative 173; Mismatches 326; Indels 192; Gaps 39;  
 QY 65 CALRPSGG-----TVYEAAYE-----VDYSASITLQVLVDAPCNISCL 104  
 DB 28 CCLTEAMPRLALLVLLMATAWHAQGVPIQSPGPELVVEFGTTVTLRCV-----GNGSVE 83  
 QY 105 W----VFKHSSLNCP-----HFDLQNRGVSMVILKMTETOAGE-----YLLFTQ 146  
 DB 84 WDGPISPHNLDLPPSSILLTNNATFQNTCTYHC-----TEPGNPRGGNATHLYVK 136  
 QY 147 SEATNYTIL---FTVSTRNTLLYTLRRPYFRKMNQDALV-CISESVPEPIVE--WVLC 200  
 DB 137 DPARPWKVLAQEVTV-----LEGQDALLPCL---LTDPALEAGVSLVR 176  
 QY 201 SOGESCKEE-----SP---AVVKEEKLHFLGTDIRCCARNELGRECTRL-----FTI 247  
 DB 177 VRGRPVLRQTNWSPFPHWGFTIHKAFIENHVY-----QCSARVD-GRVTSMGIWLKVK 231  
 QY 248 DLNQTPTTL-PQLFLKV-GEPLWIRCKAVHNHVGFLUTWELENKALBEGN---YFEMS 301  
 DB 232 DISGPATLTLEPAELVRIOGEAAQIVCSASINDVNDV-----SLRHGDTKLTIQSS 284  
 QY 302 TYSTNRTMIRILFAFVSSVARNDTGYTCCSSK---HPSQSALVTIVGKFINATFN-SSE 357  
 DB 285 DFHDNRVQ-KVLTNLNDHVSFQDAGNYSCTATNAGNHSASWFRVVSAYSLNTSEOSL 343  
 QY 358 DYEIDQVEECFVSFRKAYPOIRC-TWTFSRKSPCEQKGLD-----NGYSISKFCN--- 408  
 DB 344 LOEVTVGKVDLQVKVEAYPGLESFNWY-LGPFSDYQDKLDFVITKDYTYRTSLSPR 402  
 QY 409 -HKHQPGEYIFHAENDDAQFTKMTNLIRRKPOVLAEASASQAS-----CFSDGYPPLSWT 463  
 DB 403 LKRSEGRYSFLARNAGQONALTETLRYPEVRVMTWLINGSDTLLCEASGVPQPSVT 462  
 QY 464 WKCKSDKSPNCTEE---ITEGVNKRKANRVFGVOMSSSTLMSSEAIKGFVKKCAYNLS 520  
 DB 463 WVQCRSHTRDCEASAGLVLEDSEVLSQVPFYEIVHSLAIGTLEHNRITYECRAFSV 522  
 QY 521 GTCETILLNSPGPFPIQDNIISFYATIGVC---LLFVIVTLTLLCHYKKQFYESOLO 577  
 DB 523 GNSQTFWPISIGAHTPLPDELLETPVLLTCTMSIMALLLLLLLLLYKKQKPKYQVRWK 582  
 QY 578 MVQVTGSSDNEYFVDFREVEYDLKWEFPRENLEFGVLGSGAFGKVMNATAYGISKTV 637  
 DB 583 IIE--SYEGNSYFIDTQTPYNEKWEFPNRNLQFGKTLGTGAFGVVEATFGLGRKDA 640  
 QY 638 SIQVAVKMLKEKADSSERALEMSELKMMTQLGSHENIVNLGACTLSGPIYLIFEYCCYG 697  
 DB 641 VLKAVVAKLSTAHADKEALMSELKIMSHLGOHENIVNLGACTHGGPVLVITFEYCCYG 700  
 QY 698 DLLNLYRSKRE-----KFRWTWTEIFKEHNFSPYPTFQSHPNSSM 737  
 DB 701 DLLNFLRRQAEAMLGPSLSVQGDPEAGAGYKNIHLEKKYVRDRSGFS-----747  
 QY 738 PGREVOIHPDSQOISGLHGNSPHSEDEIEVENOKRLEEEEDNLVTFEDLLCFAYQVAK 797  
 DB 748 --SGVDTYIVEMRPVSTSSNDSEEDLGKEDGRPLE-----LRDLLHFSQVAAQ 796  
 QY 798 GMEFLEPKSVHRDLAARNVLVTHGKVVKICDFGLARDIMSDSNVYVRGNARLPVKWMAP 857  
 DB 797 GMAFLASKNCIHRDVAARNVLLTSGRVAKIGDFGLARDIMSDSNVYVRGNARLPVKWMAP 856

QY 858 ESIFEGIVTIKSDVNSYGILLWEIFSLGYNYPGIPVDANFYKLIQNGFKMDOPFYATEE 917  
 DB 857 ESIFDCVTVQSDVNSYGILLWEIFSLGYNYPGILVNSKFYKLVKDGQMAQAPAPKN 916  
 QY 918 IYIIMOSWAFDRSRKSPNLTSLGCLADAEAMYQNVDRV---SECPHYQNR-- 972  
 DB 917 IYIIMOSWAFDRSRKSPNLTSLGCLADAEAMYQNVDRV---SECPHYQNR-- 972  
 QY 973 RPFSP 977  
 DB 968 RPFSP 972

RESULT 11  
 KEMS\_HUMAN  
 ID KEMS\_HUMAN STANDARD: PRT: 972 AA.  
 AC P07333;  
 DT 01-APR-1998 (Rel. 07, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Macrophage colony stimulating factor 1 receptor precursor (CSF-1-R)  
 DE (EC 2.7.1.112) (fms proto-oncogene) (c-fms) (CD115 antigen).  
 GN CSF1R OR FMS.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89239490; PubMed=2524025;  
 RA Hampe A., Shamoon B.M., Gobet M., Sherr C.J., Galibert F.;  
 RT "Nucleotide sequence and structural organization of the human FMS  
 proto-oncogene.";  
 RL Oncogene Res. 4:9-17(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86175013; PubMed=2421165;  
 RA Coussens L., van Beveren C., Smith D., Chen E., Mitchell R.L.,  
 RA Isacke C.M., Verma I.M., Ullrich A.;  
 RT "Structural alteration of viral homologue of receptor proto-oncogene  
 fms at carboxyl terminus.";  
 RL Nature 320:277-280(1986).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97179223; PubMed=9027509;  
 RA Andre C., Hampe A., Lachaux P., Martin E., Wang X.P., Manus V.,  
 RA Hu W.X., Galibert F.;  
 RT "Sequence analysis of two genomic regions containing the KIT and the  
 FMS receptor tyrosine kinase genes.";  
 RL Genomics 39:216-226(1997).  
 RN [4]  
 RP SEQUENCE OF 1-16 FROM N.A.  
 RX MEDLINE=86261741; PubMed=2524648;  
 RA Visvader J., Verma I.M.;  
 RT "Differential transcription of exon 1 of the human c-fms gene in  
 placental trophoblasts and monocytes.";  
 RL Mol. Cell. Biol. 9:1336-1341(1989).  
 RN [5]  
 RP SEQUENCE OF 1-16 FROM N.A.  
 RX MEDLINE=86281820; PubMed=3525854;  
 RA Wheeler E.F., Roussel M.F., Hampe A., Walker M.H., Fried V.A.,  
 RA Look A.T., Rettenmier C.W., Sherr C.J.;  
 RT "The amino-terminal domain of the v-fms oncogene product includes a  
 functional signal peptide that directs synthesis of a transforming  
 glycoprotein in the absence of feline leukemia virus gag sequences.";  
 RL J. Virol. 59:224-233(1986).  
 CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN  
 CC TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-

CC PROTEIN KINASES.  
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC -!- DATABASE: NAME=PROW; NOTE=PROW 1:13-21(2000);1999017666.g.htm".  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/guide/1999017666.g.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC EMBL: M25786; AA58421.1; -  
 CC EMBL: U63963; AA51696.1; -  
 CC EMBL: M14002; AA35849.1; -  
 CC EMBL: X03663; CAA27300.1; -  
 CC PIR: S08123; TVHUMD.  
 CC HSP: P11362; LFGL.  
 CC Genes: HGNC:2433; CSF1R.  
 CC MIM: 164770; -  
 CC InterPro: IPR000719; Euk\_pkinase.  
 CC InterPro: IPR003006; Ig\_MHC.  
 CC InterPro: IPR003598; Ig\_C2.  
 CC InterPro: IPR003600; Ig\_Like.  
 CC InterPro: IPR001824; RTKaseIII.  
 CC InterPro: IPR001245; Tyr\_pkinase.  
 CC Pfam: PF00047; Ig; 4.  
 CC Pfam: PF00069; pkinase; 1.  
 CC ProDom: PD000001; Euk\_pkinase; 2.  
 CC SMART: SM00410; IG\_Like; 3.  
 CC SMART: SM00408; IGC2; 1.  
 CC SMART: SM00219; Tyrc; 1.  
 CC PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE: PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 CC PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 CC Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;  
 CC Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;  
 CC Immunoglobulin domain; Repeat; Polymorphism.  
 FT SIGNAL 1 19  
 FT CHAIN 20 972  
 FT MACROPHAGE COLONY STIMULATING FACTOR I  
 FT RECEPTOR.  
 FT DOMAIN 20 512  
 FT EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 513 537  
 FT POTENTIAL.  
 FT DOMAIN 538 972  
 FT CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 24 104  
 FT IG-LIKE C2-TYPE DOMAIN 1.  
 FT DOMAIN 107 197  
 FT IG-LIKE C2-TYPE DOMAIN 2.  
 FT DOMAIN 204 298  
 FT IG-LIKE C2-TYPE DOMAIN 3.  
 FT DOMAIN 299 399  
 FT IG-LIKE C2-TYPE DOMAIN 4.  
 FT DOMAIN 400 505  
 FT IG-LIKE C2-TYPE DOMAIN 5.  
 FT DOMAIN 582 910  
 FT PROTEIN KINASE.  
 FT NP\_BIND 588 596  
 FT ATP (BY SIMILARITY).  
 FT BINDING 616 616  
 FT BY SIMILARITY.  
 FT ACT\_SITE 778 778  
 FT DISULFID 42 84  
 FT DISULFID 127 177  
 FT DISULFID 224 278  
 FT DISULFID 419 485  
 FT MOD\_RES 699 699  
 FT PHOSPHORYLATION (AUTO-) (IN VITRO) (BY  
 FT SIMILARITY).  
 FT MOD\_RES 708 708  
 FT PHOSPHORYLATION (AUTO-) (IN VIVO) (BY  
 FT SIMILARITY).  
 FT MOD\_RES 809 809  
 FT PHOSPHORYLATION (AUTO-) (IN VITRO) (BY  
 FT SIMILARITY).  
 FT CARBOHYD 45 45  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 73 73  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 153 153  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 240 240  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 275 275  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 302 302  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 335 335  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 353 353  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).

```
FT CARBOHYD 412 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 428 428 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 480 480 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 969 969 Y -> C (IN DBSNP:1801271).
FT CONFLICT 54 54 /FTID=VAR_011953.
FT SEQUENCE 972 AA: 107983 MW: A8D99BE23757FE8 CRC64;

Query Match 23.1%; Score 1216.5; DB 1; Length 972;
Best Local Similarity 32.8%; Pred. No. 7.7e-74;
Matches 325; Conservative 156; Mismatches 316; Indels 195; Gaps 33;

QY 75 VYEEAAVEDV--SASITLQVLVDAPGNSICLWVFKHSLNCPHFQDLQNGRVVSVILK 132
DB 22 VLESPVELVPGATVILRCV-----GNGSVWDGPPS-----PHTLSDGSSSILSTN 72
QY 133 MTE-----TQAGEYL-----LFQSBATNTILFTVSIRNTLYTLRRPYFRKM 176
DB 73 NATFQNTGTYRCTEPCDPLGGSAAILHYVKDPARPWNVL-----AQEVVVF----- 118
QY 177 ENQDALV-CISESVPEPIVWVLCDSQGSCKEESPAVYKKEK-----VLHE--- 223
DB 119 EDQDALLPCL--LNDPVLE-----AGSVLRVVRGRPLMRHTNYFSFSPWHGFTIHRKF 169
QY 224 LFGTDIRCCARNELGRECTRLFTIDLNQPTTLP-----OLFLKVGEPILIRCK 273
DB 170 IQSDYQCSALMG-GR--KVMSISIRLVKQVIGPPALTLVPAELVIRGAAQIVCS 225
QY 274 AVVHNHGRGLTWELNKALEEGNFEMSTYSTRMIRILFAFVSSVARNDTGYTCSSS 333
DB 226 ASSVDVDFVLOHNTKLA---IQOSDFHNNRYQ-KVLTNLNQDVFQHAGNYSVCVAS 281
QY 334 ----KHPSQALVTIVKGFINATNSSED---YEIDQYEEFCFSVRFKAYPQIR-CTWTF 385
DB 282 NVQGRKH-STSMFRVYVESAYLNL--SSEQNLIQEVTVGEGNLKVMVEAYPGQLGFWNTY 338
QY 386 SRKSPFCQKGLDNGYSISKFNHKKHOP-----GEYIF 418
DB 339 -----LGFSDHQPEPKLANATTKDTRHHTLTSLPLPKPEAGRYSF 381
QY 419 HAENDDAQTKMTLIRKPKQVLAEASQAS----CFSDGYPLPSWTWKCKSDKSPNC 474
DB 382 LARNPGGNRALFELTLRYPPEVSVIWTFTNGSGLTLLCAASGYPPQVNIWLQCSGHTDCR 441
QY 475 TEEITEGWVN----RKANKRVQGVWVSSSTLNMSEAIKGLVKCCAYNSLGTSCETILLN 530
DB 442 DQAQVLQVMDDEPYEVLSEQEPFKHVTQVSLTAVETLEHNTQYECRAHNSVYSGSWAFIPI 501
QY 531 SPGPFPIQDNISFYATIGVC---LLFIVVLTLLCHKYKQFYESQLQMVQVTCSSDN 587
DB 502 SAGATHPPDEFLTPTPVVACMSIMALLLLLLLLLYKYKQPKYQVRWKIIE--SYEGN 559
QY 588 EYFYVDREYEDLKEWFFRENLEFGVLGSGAFGKVMNATAYGISKTGYSIQVAVKMLK 647
DB 560 SYTFIDPTQLPNEKWEFFRNPNLQFKTLGAGAFGVVEATAFGLCKEDAVLKVAVKMLK 619
QY 648 EKADSSERALMSLKMMTQLGSHENIVNLGACTLSUGIYLIPIFYCCYGDLLNLYRSKR 707
DB 620 STAHADKEALMSLKMHLGQHENIVNLGACTHGGPVLVITYECCYGDLLNLFRRKA 679
QY 708 E-----KFRHTWTIEFKHEHNSFYFTTQSHPSNSSMPGSRVQIHP 747
DB 680 EAMLGPSLPGQDPGGVDYKNHLEKKYVRDSDGS-----SQGVDTYV 724
QY 748 DSDQISGLHNSFHSDETEYENQKLEEEEDLNVLTFEDLLCFAYQVAKGMFELEFKSC 807
DB 725 EMRPVYSTNSDSF-SEQDLDKEDGRPLE-----LRDLLHFSQVQAQGMFLASKNC 774
QY 808 VHRDLAARNVLVTHGKVKVICKDFGLARDIMSDNSNYVVRGNARLPVVKWMAPELFEGIYI 867
DB 775 IHRDVAARNVLTNGHVAKIGDFGLARDIMSDNSNYVVKGNARLPVVKWMAPESFIDCVTV 834
QY 868 KSDVWSYGILLWEIFSLGVNYPFGIPVDANFYKLIQNGFKMDQPFYATETIYIIMOSCA 927
```

```
DB 835 QSDVWSYGILLWEIFSLGVNYPFGIPVDANFYKLIQNGFKMDQPFYATETIYIIMOSCA 894
QY 928 FDSKRKRPSFNLTSFLGCQL-ADAEEAMYNQV 958
DB 895 LEPTHRPTFOQICGFLQEQAEQRRDRDYNL 926

RESULT 12
KIT_CHICK
ID KIT_CHICK STANDARD; PRT: 960 AA.
AC Q08156;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mast/stem cell growth factor receptor precursor (EC 2.7.1.112) (SCFR)
DE (Proto-oncogene tyrosine-protein kinase Kit) (c-kit).
GN KIT.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn; TISSUE=Brain;
RX MEDLINE=93292995; PubMed=7685729;
RA Sasaki E., Okamura H., Chikamune T., Kanai Y., Watanabe M.,
RA Naito M., Sakurai M.;
RL "Cloning and expression of the chicken c-kit proto-oncogene.";
RL Gene 128:257-261(1993).
CC -! FUNCTION: THIS IS THE RECEPTOR FOR STEM CELL FACTOR (MAST CELL
CC GROWTH FACTOR). IT HAS A TYROSINE-PROTEIN KINASE ACTIVITY.
CC BINDING OF THE LIGANDS LEADS TO THE AUTOPHOSPHORYLATION KIT AND
CC ITS ASSOCIATION WITH SUBSTRATES SUCH AS PHOSPHATIDYLINOSITOL
CC 3-KINASE (PI3K).
CC -! CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -! SUBCELLULAR LOCATION: Type I membrane protein.
CC -! TISSUE SPECIFICITY: HIGH IN THE BRAIN AND TESTES AND ALSO
CC PRESENT IN THE BURSA OF FABRICIUS, HEART, KIDNEY, LUNG, SPLEEN
CC THYMUS AND OVARY.
CC -! SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -! SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D13225; BAA02506.1;
CC F01R; JN0677; JN0677.
CC HSP; P11362; IFGK.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR003598; Iq_C2.
CC InterPro: IPR001824; RTKinaseIII.
CC InterPro: IPR001245; Tyr_pkinase.
CC Pfam: PF00047; Iq; 3.
CC Pfam: PF00069; kinase; 1.
CC ProDom: PD000001; Euk_pkinase; 2.
CC SMART: SM00408; IGC2; 1.
CC SMART: SM00219; TyrKc; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;
CC Transferase; Glycoprotein; Phosphorylation; ATP-binding; Signal;
CC Immunoglobulin domain.
```





FT	DOMAIN	299	397	IG-LIKE C2-TYPE DOMAIN 4.
FT	DOMAIN	398	503	IG-LIKE C2-TYPE DOMAIN 5.
FT	DOMAIN	580	913	PROTEIN KINASE.
FT	NP_BIND	586	594	ATP (BY SIMILARITY).
FT	BINDING	614	614	ATP (BY SIMILARITY).
FT	ACT_SITE	776	776	BY SIMILARITY.
FT	DISULFID	42	84	POTENTIAL.
FT	DISULFID	127	177	POTENTIAL.
FT	DISULFID	224	278	POTENTIAL.
FT	DISULFID	417	483	POTENTIAL.
FT	MOD_RES	697	697	PHOSPHORYLATION (AUTO-) (IN VITRO).
FT	MOD_RES	706	706	PHOSPHORYLATION (AUTO-) (IN VIVO).
FT	MOD_RES	807	807	PHOSPHORYLATION (AUTO-) (IN VITRO).
FT	CARBOHYD	45	45	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	73	73	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	302	302	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	335	335	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	389	389	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	410	410	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	449	449	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	478	478	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	491	491	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CONFLICT	446	446	Q > E (IN REF. 1).
FT	CONFLICT	553	553	R > S (IN REF. 1).
FT	CONFLICT	616	616	L > I (IN REF. 1).
FT	CONFLICT	744	744	A > H (IN REF. 1).
FT	CONFLICT	814	814	MISSING (IN REF. 1).
FT	CONFLICT	977	977	MISSING (IN REF. 1).
FT	SEQUENCE	977	977	AA: 109178 MW; 7EDF8310CCF98906 CRC64;
QY	Query Match	22.8%;	Score 1204;	DB 1; Length 977;
DB	Best Local Similarity	33.2%;	Pred. No. 5.3e-73;	
DB	Matches 322;	Conservative 155;	Mismatches 330;	Indels 164; Gaps
QY	83	VDVSASITLQVL----	VDAPGNTSCLWFKHSSILNCOPHFDLQNRGVV----	SMVLK 133
DB	32	VEPGETVTLRCVNSGVSEWDGPISPYKTLDPES----	PGSTLTTRNATFNKWTGYRCTEL 87	
QY	134	TETOAGEYL--LFIQSEATNYTIL--	FTVYSIRNTLLYTLRRPYFRKMEQNQDALVCISES 188	
DB	88	EDPWAGSTTHLYVKDPAHSMNLLAQEVTV-----	VEGQEA-----	124
QY	189	VPEPIVFWLDCDSOGESCKEESPAVVK-----	BEKVLHELFGTDIRCCAR 234	
DB	125	LPCLITDPAALKDSY--SLMREGGRQVLRKTYFTYFSPWRGFIIRAKVL-----	DSNTYVYCKT 179	
QY	235	NELGRECTRL--FTIDLN----	QTPQTL--POLFLKV--GEPLWLRCKAVHVNHGFGTLWEL 267	
DB	180	MVNGRESTGTGWLKVRNVRHVEPPQIKLEPSKLVIRGEAAQIVCSATNAEVEGNVLKR 239		
QY	288	ENKALE-----	BGNTFEMSTYSTNRTMIRILFAFVYSVARNDTGYTTC--SSSKHPQS 339	
DB	240	GDTKLEIPLNSDFQDNYK-----	KVRALSUNAVDFQDAGIYSCVASNDVGTRT 268	
QY	340	ALVT--IVGKGFINATN--SEDEYIDQYEBEFCFSVRKPAIQPIR--CTWTSFRKSFPCEQK 395		
DB	289	ATMNFQVESAYLNLTSQESLILQEVSGDSILITVHADAYPSIQHYNTVTLGPFPE--DQR 347		
QY	396	GLD-----	NGYSISKFCNH--KHOPGEYIFHAENDDAQFTKMTFLNIRKKPOVLAEAS 446	
DB	348	KLEIFTORAIRYRTFKLFLNRVRKASEAGQYFLMAQNKAGWNNTFELTLRYPPEVSVYTM 407		
QY	447	ASQAS----	CFSDGPPPLPSMTWKKCDSPKSPNCTEITEGVYN-----RKANRKYFGQWVSS 498	
DB	408	PVNGSDVLFCDYSGYPOPSVYTWMECRGHTDRCDQAQALQVWNTDHEVLVSQKFPDKVIO 467		
QY	499	STLNMSEAIKGLVKCCAYNSLGTSCETILLNSGPGPFPIQDNISFYATIGVC---LLFI 555		
DB	468	SOLPIGTLKHNMTVEFKTHNSVGNSSQYFRAVSLGQSKQLPDESLETFPVVVACMSVMSLL 527		
QY	556	VVLTLILCHYKKOFRYESQLQWQVTVGSSDNEYFYVDFREYEDLKWEPRENLEFGVK 615		
DB	528	VLLLLLLLYKKQPKQYQVARKWIIIE--RYEGNSYTFIDPTQLPYNEKWEPPRNLOFGKT 585		

QY 616 LGSAGFKVMNATAYGISTGYSIOVAVKMLREKADSSERREALMSKMTOLGSHENIV 675  
 DR ||:|||||: |||:|: ||:|||||: ||:|:|||||:|:|: ||:|||||  
 Db 586 LGAGAFGKVKVATEAGLGKEDAVLVAVKMLAKSTAHADKEALMSKIMSHLGQHENIV 645  
 DR ||:|||||: |||:|: ||:|||||: ||:|:|||||:|:|: ||:|||||  
 QY 676 NLLGACTLGGPIVLYFECYCGDLLNLYRSKRE-----KPHRTWT 715  
 DR ||:|||||: |||:|: ||:|||||: ||:|:|||||:|:|: ||:|||||  
 Db 646 NLLGACTHGGPVLVITEYCCYCGDLLNLYRSKRE-----KPHRTWT 705  
 DR ||:|||||: |||:|: ||:|||||: ||:|:|||||:|:|: ||:|||||  
 QY 716 ELFKHEHNSFYPTTFOHSPNSMPGSGREVQIHDPDSQISGLHNSHSEDEIEYENOKRLE 775  
 DR ||:|||||: |||:|: ||:|||||: ||:|:|||||:|:|: ||:|||||  
 Db 706 YVRDSGFS-----SQGVDTYVEMRPVSTSSDSFFKQD-LKEASRPLE 749  
 DR ||:|||||: |||:|: ||:|||||: ||:|:|||||:|:|: ||:|||||  
 QY 776 EEDLNVLTFEDLLCFAYQVAKGMEFLFKSCVHRDLAARNVLVTHGKVVKICDFGLARD 835  
 DR ||:|||||: |||:|: ||:|||||: ||:|:|||||:|:|: ||:|||||  
 Db 750 -----LWDLHFSSQVAGMAFLASKNCIHRDVAARNVLLTSGHVAKIGDFGLARD 800  
 DR ||:|||||: |||:|: ||:|||||: ||:|:|||||:|:|: ||:|||||  
 QY 836 IMSDSNVYVRGNARLPVKWAPESLFEIGYTIKSDVWSYGILLWEIFSLGVNYPGPVD 895  
 DR ||:|||||: |||:|: ||:|||||: ||:|:|||||:|:|: ||:|||||  
 Db 801 IMSDSNVYVRGNARLPVKWAPESLFEIGYTIKSDVWSYGILLWEIFSLGVNYPGPILVN 860  
 DR ||:|||||: |||:|: ||:|||||: ||:|:|||||:|:|: ||:|||||  
 QY 896 ANFYKLQNGKMDQFPVATEEIIYIMQSWAFDSKRPSPNLTSLFCOLADAEEAMY 955  
 DR ||:|||||: |||:|: ||:|||||: ||:|:|||||:|:|: ||:|||||  
 Db 861 NKFYKLVKDGVMQAPVFPAPKNIYSIMOSWDLPTTRPTFOQICFL-----QEQARL 914  
 DR ||:|||||: |||:|: ||:|||||: ||:|:|||||:|:|: ||:|||||  
 QY 956 QNVGGRVSECP 966  
 DR ||:|||||: |||:|: ||:|||||: ||:|:|||||:|:|: ||:|||||  
 Db 915 ERRDQDYANLP 925  
 DR ||:|||||: |||:|: ||:|||||: ||:|:|||||:|:|: ||:|||||

RESULT 14  
 KFMS\_RAT STANDARD; PRT; 978 AA.  
 ID KFMS\_RAT AC Q00495;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Macrophage colony stimulating factor I receptor precursor (CSF-1-R)  
 DE (EC 2.7.1.112) (fms proto-oncogene) (c-fms).  
 GN CSF1R OR CSFMR OR FMS.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar; TISSUE=Skeletal muscle;  
 RX MEDLINE=93001225; PubMed=1389227;  
 RA Borycki A.G.; Guiller M.; Leibovitch M.P.; Leibovitch S.A.;  
 RT "Molecular cloning of CSF-1 receptor from rat myoblasts. Sequence  
 analysis and regulation during myogenesis.";  
 RL Growth Factors 6:209-218(1992).  
 CC -!- FUNCTION: THIS PROTEIN IS THE RECEPTOR FOR CSF-1, IT IS A PROTEIN  
 CC TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; X61479; CAA43706.1;  
 DR PIR; S16385; S16385.  
 DR HSSP; P11362; 1FGK.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR InterPro; IPR003006; Ig\_MHC.

DR InterPro; IPR003598; Ig\_C2.  
 DR InterPro; IPR003600; Ig\_Like.  
 DR InterPro; IPR001824; RTKinaseIII.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00047; Ig; 4.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 2.  
 DR SMART; SM00410; IG\_Like; 3.  
 DR SMART; SM00408; IGC2; 1.  
 DR SMART; SM00219; TyrKC; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 KW Proto-oncogene; Tyrosine-protein kinase; Receptor; Transmembrane;  
 KW Transferrase; Glycoprotein; Phosphorylation; ATP-binding; Signal;  
 KW Immunoglobulin domain; Repeat.  
 FT SIGNAL 1 19 BY SIMILARITY.  
 FT CHAIN 20 978 MACROPHAGE COLONY STIMULATING FACTOR I  
 FT RECEPTOR.  
 FT DOMAIN 20 511 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 512 536 POTENTIAL.  
 FT DOMAIN 537 978 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 24 104 IG-LIKE C2-TYPE DOMAIN 1.  
 FT DOMAIN 107 197 IG-LIKE C2-TYPE DOMAIN 2.  
 FT DOMAIN 204 298 IG-LIKE C2-TYPE DOMAIN 3.  
 FT DOMAIN 299 397 IG-LIKE C2-TYPE DOMAIN 4.  
 FT DOMAIN 398 503 IG-LIKE C2-TYPE DOMAIN 5.  
 FT DOMAIN 580 914 PROTEIN KINASE.  
 FT NP\_BIND 586 594 ATP (BY SIMILARITY).  
 FT BINDING 614 614 ATP (BY SIMILARITY).  
 FT ACT\_SITE 776 776 BY SIMILARITY.  
 FT DISULFID 42 84 POTENTIAL.  
 FT DISULFID 127 177 POTENTIAL.  
 FT DISULFID 224 278 POTENTIAL.  
 FT DISULFID 417 483 POTENTIAL.  
 FT MOD\_RES 697 697 PHOSPHORYLATION (AUTO-) (IN VITRO)  
 FT MOD\_RES 706 706 PHOSPHORYLATION (AUTO-) (IN VIVO)  
 FT MOD\_RES 807 807 PHOSPHORYLATION (AUTO-) (IN VITRO)  
 FT CARBOHYD 45 45 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 335 335 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 410 410 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 449 449 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 478 478 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 491 491 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 978 AA; 109264 MW; 0A68456EF56BC7E3 CRC64;  
 Query Match 22.6%; Score 1194; DB 1; Length 978;  
 Best Local Similarity 33.1%; Pred. No. 2.5e-72;  
 Matches 321; Conservative 154; Mismatches 332; Indels 164; Gaps 32;  
 QY 83 VDVSASTLQVL-----VDAPGNISCLMVFKHSSLCNCPHFDLQNRGV-----SWILKM 133  
 Db 32 VEPGETVTLRCVSGSVGWDGPISPYWTLDPE-----PGSTLTTRNATFKNTGYRCTL 87  
 QY 134 TETQAGEYL-LFTQSEATNYTL--FTVSIKNTLLYTLRRPYFRKMENODALVCISES 188  
 Db 88 EDPMAGSTTHLYVKDPAHNSNLLAQEVTV-----VEGQEV----- 124  
 QY 189 VPEPIVEMWLCDSDGSCKEESPAAVKK-----EKKVLHELFGTDIRCCAR 234  
 Db 125 LPLITDPALKDSV-SLMREGROVLKTYVYFSAWRGFIIRKAKVL-----DSNTYVCKT 179  
 QY 235 NEIGRECTRL-FTIDLN-----QTPTTL-PQLFLKV-GEPLMIRCKAVVHNGFGLTWEL 287  
 Db 180 MNGRESTSTGILKLVNRVHPEPPQIKLEPSKLVIRGEAAQIVCSATNAEVGFNVILKR 239

QY 288 ENKALE-----EGNYFEMSTYSTRNTRILFAFVSSVARNDTGYTC-SSSKHPSQS 339  
 Db 240 GOTKLEIPLNSDFQDNYK-----KYRALSINAVDFODAGIYSCVASNDVGTRT 288  
 QY 340 ALVT--IVGKGFINATN--SSEYIDIOYEFECFVRKAYPQIR-CTWTFSRKSFCEQK 395  
 Db 289 AFMNFQVVEAYLNLTSQSLQEVSGDSLITVHADAYPSIQHYNWTVLGYPFFE-DQR 347  
 QY 396 GLD-----NGYSISKFCNH--KHQPGYIIPHAENDDAQFTKMTFLNTRRKQVLAAS 446  
 Db 348 KLEFTQRIAYITKFLFLNRVKASEAGYFLMAONKAGWNNTFTLTLYRPEVSVTW 407  
 QY 447 ASQAS-----CFSDGYPPLPSWTWKCKDKSPNCTEITEGVN----RKANRKFVGOWSS 498  
 Db 408 PYNGSDVLFCDVSGYPPQSVTWMECHRGHTRDCDEAALQVWNTDHPVLSQKFFDKVIIQ 467  
 QY 499 STLNSEAIKGLVKKCAVNSLGTSCETILLNSPGPFPIQDNISYATIGVC---LLFI 555  
 Db 468 SOLPIGTLKHNMTYFCKTHNSVGNSSQYFRAVSLGQSKQLPDBESLFTPVVACMSVMSLL 527  
 QY 556 VVTLTLLCHKKYKOFYESQLOMVOVGTSSDNEYFYVDREYEDYDKWEPFRENLEFGK 615  
 Db 528 VULLLLLYKKYKOPKYQVRWKIIE--RYEGNSYTFIDPTQLPYNKKEWFFPRNNLOFGKT 585  
 QY 616 LGSGAFGKVMNATAYGISKTVGSIOAVKMLKEKADSSEREALMSLKMNTOLGSHENIV 675  
 Db 586 LGAGAFGKVVETATAGLGEDAVLKVAVKMLKSTAHADKEALMSLKMNTOLGSHENIV 645  
 QY 676 NLGACTLGGPIYLFIFCYCYGDLNLNLSKRE-----KFHRTWT 715  
 Db 646 NLGACTHGGPVLVITEYCCYCGDLNLFRLKRAEAMGLPSLQSDSEGSSYKNIHLEKK 705  
 QY 716 EFKENHSFYTFQSHPNSSMPGSRVQIHDPDSDDOISGLHNSFSEDEIEYENOKRLE 775  
 Db 706 VYRDSGFS-----SGQVDTYVEMRPVSTSSDSDFKQD-LDKEPSRPLE 749  
 QY 776 EEDENLVTFFDLFCAYOVAKGMEFLFKSCVHRDLAARNVLVTHGVKVKICDFGLARD 835  
 Db 750 -----LWDLHSSQVAGMAFLASKNCIHRDVAARNVLLTSGHVAKIGDFGLARD 800  
 QY 836 IMDSNVYVGNARLPVKWAPESLFEIYITKSDVMSYGILLWEIFSLGVNPPYGPVD 895  
 Db 801 IMDSNVYVGNARLPVKWAPESILYCYVTQSDVMSYGILLWEIFSLGVNPPYGPILVN 860  
 QY 896 ANFYKLIQNGFKMDQPFYATEIYIIMQSCWAFDSKRFSFNLTSFLGQLADAEAMY 955  
 Db 861 NKFKYLVKDGQYMAQPFVAFKPIYISIMQSCWDLPTRRPTFOQICELL-----QEQARL 914  
 QY 956 QNVDCRVSECP 966  
 Db 915 ERRODYANLP 925

RESULT 15  
 PGDS\_RAT  
 ID PGDS\_RAT STANDARD; PRT: 1088 AA.  
 AC P20786;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Alpha platelet-derived growth factor receptor precursor (EC 2.7.1.112)  
 DE (PDGF-R-alpha).  
 GN PDGFRA.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_taxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley;  
 RA MEDLINE=90220609; PubMed=2157969;  
 RX Lee K.H., Bowen-Pope D.F., Reed R.R.;  
 RT "Isolation and characterization of the alpha platelet-derived growth

factor receptor from rat olfactory epithelium.";  
 Mol. Cell. Biol. 10:2237-2246(1990).  
 [2]  
 RN SEQUENCE OF 33-524 FROM N.A.  
 RP MEDLINE=93305723; PubMed=8318539;  
 RX Herren B., Weyer K.A., Rouge M., Loetscher P., Pech M.;  
 RA "Conservation in sequence and affinity of human and rodent PDGF  
 RT ligands and receptors.";  
 RL Biochim. Biophys. Acta 1173:294-302(1993).  
 CC -!- FUNCTION: THIS RECEPTOR BINDS PLATELET-DERIVED GROWTH FACTOR AND  
 CC HAS A TYROSINE-PROTEIN KINASE ACTIVITY. THIS RECEPTOR CAN BIND  
 CC EITHER PDGF-A OR PDGF-B.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -!- SUBUNIT: DIMER OF EITHER ALPHA-ALPHA, BETA-BETA OR ALPHA-BETA  
 CC SUBUNITS.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -!- SIMILARITY: CONTAINS 5 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to licensed@isb-sib.ch).  
 CC  
 CC EMBL; M63837; AAA40743.1; ALT\_INIT.  
 CC EMBL; Z14118; CAA78488.1;  
 CC PIR; A34710; PRTCA.  
 CC HSP; P11362; IFGK.  
 CC InterPro: IPR000719; Euk\_pkinase.  
 CC InterPro: IPR003599; Ig.  
 CC InterPro: IPR003006; Ig\_MHC.  
 CC InterPro: IPR001824; RTKinaseIII.  
 CC InterPro: IPR001245; Tyr\_pkinase.  
 CC Pfam: PF00047; Ig; 4.  
 CC Pfam: PF00069; pkinase; 1.  
 CC ProDom: PD000001; Euk\_pkinase; 2.  
 CC SMART; SM00409; IG; 3.  
 CC SMART; SM00219; Tyrc; 1.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 CC PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 CC Tyrosine-protein kinase; Receptor; Transmembrane; Glycoprotein;  
 CC Transferase; Phosphorylation; ATP-binding; Immunoglobulin domain;  
 CC Signal.  
 CC SIGNAL  
 CC CHAIN 1 23  
 CC 24 1088 ALPHA PLATELET-DERIVED GROWTH FACTOR  
 CC RECEPTOR.  
 CC DOMAIN 24 524  
 CC TRANSMEM 525 548  
 CC DOMAIN 549 1088 EXTRACELLULAR (POTENTIAL).  
 CC DOMAIN 592 953 CYTOPLASMIC (POTENTIAL).  
 CC NP\_BIND 598 606 PROTEIN KINASE.  
 CC BINDING 626 626 ATP (BY SIMILARITY).  
 CC ACT\_SITE 817 817 ATP (BY SIMILARITY).  
 CC MOD\_RES 848 848 BY SIMILARITY.  
 CC CARBOHYD 41 41 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 CC CARBOHYD 75 75 O-LINKED (POTENTIAL).  
 CC CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 358 358 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CARBOHYD 467 467 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CONFLICT 150 150 L -> R (IN REF. 2).  
 CC CONFLICT 519 519 S -> T (IN REF. 2).  
 CC SEQUENCE 1088 AA; 122641 MW; 590C8BB0418601E7 CRC64;

```
Query Match      22.3%; Score 1176; DB 1; Length 1088;
Best Local Similarity 35.6%; Pred. No. 4.6e-71;
Matches 294; Conservative 128; Mismatches 288; Indels 115; Gaps 24;

Qy 246 TIDLNQTPQTTLPQLFLKVGELWIRKAVHVNHGFLTW-----ELENKA-----LEEGNYF 298
Db 210 TSELNLEMDTR--QVYKAGETIVVTC-AVFNEVVDLQWTPYGEVRNKGITMLEE-----262

Qy 299 EMSTYSTNRTWIRILFAF-VSSVARNDGYTCS---SSKHPQSALVTIV--GKGFIN- 351
Db 263 -----IKLPSIKLIVYTTVTKATVKDSDGECAARQATKEVKEMKTVTISVHREGFVOI 316

Qy 352 --ATNSSEDEYIDQYEEFCFSVRKAYPQIRCTWTFSRKSF-----PCQKGLDNGY 401
Db 317 RPTFGHLETVNLHVREFV--VEVQAYTPPRISWLKDNLTJENLTETITDVQORSQETRY 374

Qy 402 -SISKFCNHHQO-PGEYIFHAENDDAQFTKMETLNIIRKQPVL-----AEASASQASC 452
Db 375 OSKLLIRAKEDSGHYTIIVQNDDDMKSYTFELSTLPASILELVDDHHGSGGGQTVC 434

Qy 453 FSDGYPLPSWTH-----KKCSDK-----SPNCTEEITEGVNRRKANKRVFGQWVSSST 500
Db 435 TAEGTPLPNIEWMICKDKKCNNDTSWTVLASVNSNIITE--FHGRGSTVEGR-----486

Qy 501 LNMSEAIGKFLVKCCAYNSLGTSCETILINSPGPPFIQDNISFYATIGVCLIFIVVLT 560
Db 487 VSFKAVETIAVRCLAKNDLIGNRELKLVAPS----LRSELTVAARVLLVIVIVSLI 542

Qy 561 LICHYKKQFRYESQLQVQVVGSSDNEYVYDFREYEDLKWEPPRENLEFGKVLGSGA 620
Db 543 VLVVIWKQKPRYIEIRWVIESIPDGHEVIYVDPQLPYDSRWEPFRDGLVLGRILGSGA 602

Qy 621 FGVVNATAYGISKTVSIOVAVMKLKEKADSSEREAALMSELKMMTQLGSHENIVNLLGA 680
Db 603 FGVVVEGTAYGLSRQPVKAVKMLKPTARSEKQALMSELKIMTHLGPLNIVNLLGA 662

Qy 681 CTLGPIYLIFCYCYGDLNLYLSKREKF-----HRTWTEI 717
Db 663 CTKSGPIYIITEYCYGDLNLYLHKNRDSFMRHPEKPKKDDIFGLNPADESTRSYVIL 722

Qy 718 FKEHNFSEYPTFQSHPNSSMPGSRVQIHPDSDOIISGLHGN--SPHSEDEIEYENOKRLE 775
Db 723 SFENNGDYVDMKQADTQYVPMLEKREKESYSDIQRSLYDRPASYYKKKSMLDSE-AKNLL 781

Qy 776 EEDLNVLTFEDLLCFAYQVAKGMEFLEFKSCVHRDLAARNVLVTHGVKVKICDFGLARD 835
Db 782 SDDSEGLTLLDLLSFTYQVARGMEFLASKNCVHRDLAARNVLLAQGKIVKICDFGLARD 841

Qy 836 IMSDSNYVVRGNARLPVKWMAPESLFEGYITIKSDVWSYGILLWEIFSLGVNYPGIPVD 895
Db 842 IMHDSNYVSKGSTFLPVKWMAPESIFDNYTLSDVWSYGVLLWEIFSLGTPYPCWAVD 901

Qy 896 ANFYKLIONGFKMDQFFYATEEYIIIMQSCWAFDSKRKPSFPNLTSLGCQLADAEAMY 955
Db 902 STFYNNKIKSGYRMAKPDHATSEVYIEMVQCNWSEPEKRPFSFVHLSIEIVENLLPGQYKKS 961

Qy 956 QNV--DGRVSECPH-----TYQNRPFPSREMDLGL 983
Db 962 EKIHLDLFLKSDHPAVARMRVSDNAYIGVTKNEEDKLEWEGGL 1006
```

Search completed: May 27, 2003, 14:32:21  
Job time : 29.0091 secs

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 14:26:24 ; Search time 71.0358 Seconds

(without alignments)  
2880.307 Million cell updates/sec

Title: US-09-919-408-4

Perfect score: 5274

Sequence: 1 MPALARDAGTVPLLVFSAM.....PFSREMDLGLSPQAQVEDS 993

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp.archaea.\*
- 2: sp.bacteria.\*
- 3: sp.fungi.\*
- 4: sp.human.\*
- 5: sp.invertebrate.\*
- 6: sp.mammal.\*
- 7: sp.mhc.\*
- 8: sp.organelle.\*
- 9: sp.phage.\*
- 10: sp.plant.\*
- 11: sp.rodent.\*
- 12: sp.virus.\*
- 13: sp.vertebrate.\*
- 14: sp.unclassified.\*
- 15: sp.rviro.\*
- 16: sp.bacteriap.\*
- 17: sp.archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1271.5	24.1	976	13 Q9W755	Q9W755 brachydanio
2	1265.5	24.0	974	11 Q63702	Q63702 rattus ratt
3	1263.5	24.0	978	11 Q63116	Q63116 rattus norv
4	1262.5	23.9	964	6 Q97744	Q97744 sus scrofa
5	1261.5	23.9	964	6 Q97Q01	Q97Q01 sus scrofa
6	1261.5	23.9	964	6 Q97Q00	Q97Q00 sus scrofa
7	1250.5	23.7	979	6 Q8WN23	Q8WN23 canis fami
8	1247	23.6	972	4 Q99662	Q99662 homo sapien
9	1244	23.6	978	6 Q9XS93	Q9XS93 canis fami
10	1230.5	23.3	977	13 Q98SU1	Q98SU1 danio nigro
11	1229	23.3	945	6 Q77589	Q77589 equus cabal
12	1228	23.3	974	13 Q98SU3	Q98SU3 danio dangi
13	1226	23.2	954	13 Q91909	Q91909 xenopus lae
14	1223.5	23.2	977	13 Q918N6	Q918N6 brachydanio
15	1221	23.2	948	6 Q9TDT7	Q9TDT7 trichosurus
16	1219	23.1	724	6 Q9MYN0	Q9MYN0 bos taurus

17	1214	23.0	977	13 Q98SU2	Q98SU2 danio kerri
18	1213	23.0	975	13 P79750	P79750 fugu rubrip
19	1211	23.0	977	13 Q98SU4	Q98SU4 danio albol
20	1141.5	21.6	1059	13 Q9DE49	Q9DE49 brachydanio
21	1126	21.4	1087	13 Q9PUF6	Q9PUF6 gallus gall
22	1111.5	21.1	1097	11 Q8R406	Q8R406 rattus norv
23	1080	20.5	1048	13 P79749	P79749 fugu rubrip
24	1079	20.5	1019	13 Q8UVR8	Q8UVR8 fugu rubrip
25	1046	19.8	923	6 Q97745	Q97745 sus scrofa
26	1038	19.7	986	13 Q8UVR9	Q8UVR9 fugu rubrip
27	965.5	18.3	563	11 Q925F7	Q925F7 rattus norv
28	952.5	18.1	1173	13 Q9PTL0	Q9PTL0 brachydanio
29	950.5	18.0	1301	13 Q8UUV9	Q8UUV9 brachydanio
30	948.5	18.0	1327	13 Q8OHL3	Q8OHL3 gallus gall
31	934	17.7	1345	11 Q8VCD0	Q8VCD0 mus musculu
32	932	17.7	1379	13 P79701	P79701 coturnix co
33	916.5	17.4	323	11 Q9EQ22	Q9EQ22 rattus norv
34	909.5	17.2	1363	11 Q91ZT1	Q91ZT1 rattus norv
35	892.5	16.9	323	11 Q9EQ24	Q9EQ24 rattus norv
36	850	16.1	819	4 Q96KM0	Q96KM0 homo sapien
37	842	16.0	766	4 Q96KM2	Q96KM2 homo sapien
38	842	16.0	785	4 Q96KM1	Q96KM1 homo sapien
39	841	15.9	824	13 Q90749	Q90749 gallus gall
40	837.5	15.9	820	4 Q96KL9	Q96KL9 homo sapien
41	836	15.9	782	6 Q9TT23	Q9TT23 oryctolagus
42	835	15.8	350	13 Q91416	Q91416 xenopus lae
43	832.5	15.8	824	13 Q91286	Q91286 pleurodeles
44	828.5	15.7	822	4 Q14672	Q14672 homo sapien
45	827	15.7	796	13 Q91287	Q91287 pleurodeles

#### ALIGNMENTS

RESULT 1

Q9W755	PRELIMINARY;	PRT;	976 AA.
ID	Q9W755	PRELIMINARY;	PRT;
AC	Q9W755;		
DT	01-NOV-1999 (Tremblrel. 12, Created)		
DT	01-NOV-1999 (Tremblrel. 12, Last sequence update)		
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)		
DE	Kit receptor tyrosine kinase.		
GN	KIT.		
OS	Brachydanio rerio (Zebrafish) (Zebra danio).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;		
OC	Cyprinidae; Danio.		
OX	NCBI_TaxID=7955;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE-99396707; PubMed-10393121;		
RA	Parichy D.M., Rawls J.F., Pratt S.J., Whitfield T.T., Johnson S.L.;		
RT	"Zebrafish sparse corresponds to an orthologue of c-kit and is		
RT	required for the morphogenesis of a subpopulation of melanocytes, but		
RT	is not essential for hematopoiesis or primordial germ cell		
RT	development."		
RL	Development 126:3425-3436(1999).		
CC	-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).		
CC	-!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-		
CC	PROTEIN KINASES.		
DR	EMBL: AF153446; AAD41890.1; -		
DR	HSSP: P11362; IFGK.		
DR	ZFIN: ZDB-GENE-980526-464; kit.		
DR	InterPro: IPR000719; Euk_pkinase.		
DR	InterPro: IPR003598; Ig_C2		
DR	InterPro: IPR003600; Ig_like.		
DR	InterPro: IPR003006; Ig_MHC.		
DR	InterPro: IPR001824; RTKinaseIII.		
DR	InterPro: IPR001245; Tyr_pkinase.		
DR	Pfam: PF00047; ig; 3		
DR	Pfam: PF00069; pkinase; 1.		
DR	ProDom: PD000001; Euk_pkinase; 2.		
DR	SMART: SM00408; IGC2; 1.		

DR SMART: SM00410; IG-like; 1.  
DR SMART: SM00219; Tyrc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00101; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
DR ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
KW Phosphorylation; Receptor; Transferase; Transmembrane;  
KW Tyrosine-protein kinase  
SQ SEQUENCE 976 AA; 109278 MW; A80AA01658C1A2DA CRC64;  
  
Query Match 24.1%; Score 1271.5; DB 13; Length 976;  
Best Local Similarity 38.5%; Pred. No. 3.4e-96;  
Matches 294; Conservative 122; Mismatches 262; Indels 85; Gaps 20;  
  
QY 253 PQTLLPQ---LFLKVGPELWIRKAVVNHGFLTWLE---NKALEEGNYFEMSYSTN 306  
DB 206 PPTLGPQORVLLTQGEKLSGSTNSVNSDIKAPKAPNGVNPVSHQNSHLLTEPIHV 265  
QY 307 RTMIRLFAFVSSVARNDTGYTCSSSKHPSQA---LVTIVGKGFNATN--SSEDEYID 362  
DB 266 RTALLSL---SSVTQDAGNYCEAINEKGTAKPVWVNIYKGFNITSDVNSTRVR 321  
QY 363 QYEEFCFSVRKAYPQIRC--TWTFSRKSFCEQKLDNGYSISKFCNH----- 409  
DB 322 AGESLSLRVVMNAYPKPHTFSWYS-----GVKLTNTDHTVITSRTHGNSYT 368  
QY 410 -----KHQGEYIFHAENDDAQFTMTLNIRKPKQVLAESA--SQACFSFGY 458  
DB 369 SELKLVRLKVSIGIYVTFSCNLRDATIROTFEVHISKQPIVSVEGIDGQVRCVAGY 428  
QY 459 LPSWTW-----KKCDKSPNCTEETEGVWNRKANRKFVGOWSSSTNMSAIGKFL 511  
DB 429 TPQIKWYCDLPHSRCSNLI--NATQE--EEDVTVTWNTPFGKAVESRLNTHK--NNYA 484  
QY 512 VKCAVNSLCTSETILLNSPGPPP---FIQDNISFYATIGVCLLFIIVTLTLCHKYKK 568  
DB 485 TLECVASANEIYVTLFISENTVPHELPLLLIGFVAAAVILVILVLT---KYMQ 540  
QY 569 QFYESQLOMVQVGTSSDNEYFYVDFREYEDLKWEPRENLEFGKVLGSGAFKVMNAT 628  
DB 541 KPKYQIQWKVIE--GIHGNVYVIDPTQLPYDHOWEPFROKLFCKTLGSGAFKGVVEAT 598  
QY 629 AYGTSKGVSIQVAKMKKEKADSSEREAALMSLKMTQLGSHENIYNLLGACTLSGPIY 688  
DB 599 AYGMSKADTVMTVAVKMLKPSAHATEKALMSLEKVLISYLGHNHINYNLLGACTVGGPTL 658  
QY 689 LIFCYCCYGLLNLVRSKRKFKHRTWTEIFEKHNFSFYPTFQSPHNSMPGSEVQ---- 744  
DB 659 VITEYCCGDLNLFRRRYVYYT---TLGEDAYRNVMMQSEPNDSRNGYMTMKPSVL 715  
QY 745 -IHPDSQISGLHGNSTHSEDEIYENQKLEEEEDNLVTFEDLLCFAYQVAKGMEFLE 803  
DB 716 GILSSERNRLSKGDSYSDSAV-----SETLQEDGLTLD--EDLLSFSYQVAKGMDFLA 769  
QY 804 FKSCVHRDLAARNVLVTHGKVKVICDFGLARDIMSDSNVYVGRNARLPVKWMAPESEFEG 863  
DB 770 SKNCIHRDLAARNILLTQGRVAKICDFGLARDITDSNVYVGRNARLPVKWMAPESEFEC 829  
QY 864 IYTKSDVWSYGILLWEIFSLGVPNPYGPVDPANFYKLIQNGKMDQPFYATEEYIIMQ 923  
DB 830 VYTFESDWSYGILLWEIFSLGSPYPCMPVDSKPYKMKIEGKYMESPEFSPEMYDIMH 889  
QY 924 SCWAFDSKRPSFNLNLFSLCCQLADAEAMQYVNDGRVSECP 966  
DB 890 SCWDADVPKRPFSKIVEKIEQQISDSTKHLYLNFSSRLPAAP 932  
  
RESULT 2  
Q63702 PRELIMINARY; PRT; 974 AA.  
ID Q63702  
AC Q63702;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE C-kit receptor tyrosine kinase isoform.  
OS Rattus rattus (Black rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_Taxid=10117;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BN/FMAI;  
RA Tsujimura T., Tono T., Yamazaki M., Nomura S., Kitamura Y.;  
RT "Two isoforms of rat c-kit receptor tyrosine kinase.";  
RL Nucleic Acids Res. 0:0-0(0).  
CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
CC -|- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
CC PROTEIN KINASES.  
DR EMBL; X62491; CAA44354.1; -;  
DR HSSP; P11362; IFGK.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR003598; Ig\_C2.  
DR InterPro; IPR003600; Ig\_Like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR001824; RTKinaseIII.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00047; Ig; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 2.  
DR SMART; SM00408; IGC2; 1.  
DR SMART; SM00410; IG\_Like; 3.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00101; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
KW Phosphorylation; Receptor; Transferase; Transmembrane;  
KW Tyrosine-protein kinase.  
SQ SEQUENCE 974 AA; 108955 MW; B9BFAD9BA0DA190D CRC64;  
  
Query Match 24.0%; Score 1265.5; DB 11; Length 974;  
Best Local Similarity 34.0%; Pred. No. 1.1e-95;  
Matches 334; Conservative 171; Mismatches 361; Indels 117; Gaps 37;  
  
QY 35 CVLINHKNDSVSGSSYPMVSESPDLGALRQSSGTVYEEAAVVDVSASTLQVL 94  
DB 12 CVLLVLLRGQGTGTSQSPASGEPSP-----SIQPAQS-----ELIVEAGDTIRLT 57  
QY 95 VDAPGNISCLVFKVHSSLNCPHFDLQNRGVVSVMLKMT--TQAGEYLLFTQSEATNYT 153  
DB 58 CTDPAFYK--WTFE-----ILDVRIENKQSEWIREKAEATHGTKYTCVSGSLRSSI 107  
QY 154 ILFTVSTRN--TLVTLRRPYFRKMNODALVCISESVPEIPEVWVLCDSQGESCKEESPA 212  
DB 108 YVF---VHDPVAVFLVGLPLFGK--EDNALVRCPLTDQ--VSNYSLIECDGKSLFDLKF 162  
QY 213 V-----VKKEKVLHLEFGTDIRCCARNELGECRLFTID-----LNQTPQTTLPOL 260  
DB 163 VPKAGITIRKVRAYHRLC---IRCAAQREGKWMRSKDKFTLKVRAAIKAIPVVSVPET 219  
QY 261 --FLKVGPELWIRKAVVNHGFLTW-----ELENKALEEGNYFEMS--TYSTNRTMIR 311  
DB 220 SHLLKEGDTFTVICTIKDVTSVDSMTIKLNPQPSKQAVKRNHSHQGDNFYERQETLT- 278  
QY 312 ILFAFVSVARNDTGYTCSSSKHPSQSALVT---IVGKGFN---ATNNSSEDEYIDQVE 365  
DB 279 -----ISSARVNDSCVFCYANNFGSANVTTLKVVKEGFNIIPVNTT--VFVTDCE 331  
QY 366 EFCFSVREKAYPQ--IRCTWTFSRKSFCEQKGLD-----NGYSISKFNHKK-----QP 413  
DB 332 NVDLVVEFEAYPKPEHQQWIYMNRT--PTNRGDEVKSDNQSNIRYVNELRLTLRLKTEG 389  
QY 414 GEYIFHAENDDAQFTMTLNIRKPKQVLA--EASASQASCFSDGYPPLPSWTWKCKSDKS 471

Db 390 GTYTFVLSNDSVSASVTEDVYVNTKPEILTYDRLMNGRLQCVAAAGPPEPTIDWYFCTGAE 449  
 QY 472 PNCETEITE-GWNNRANKRVKGVQWVSSSTLNKSEAIKGLFLVKCCAYNSLGTSCETILLN 530  
 Db 450 QRCVTVPVPPVDQIQNASVSPGKLVQSSIDSSVFRHNGTVECKASNAVGS--SAFFN 507  
 QY 531 SPGPFPF---IQDN-----ISFYATIGVCLLFIVLTLTLLCHIKYKKQFRYESQLQWY- 579  
 Db 508 ---FAKKEQIQPHLFTPTLLIGFVVVTAG--LMGIIVMWL--AYKYLOKPMYEVQWKVVE 559  
 QY 580 QVTGSSDNEYFYVDREYEDYDLKWEPPRENLEFGKVLGSGAFKGVNNAATAYGISKTGYSI 639  
 Db 560 BING---NNYVYIDPTQLPYDTHKWEPPRNRLSFGKTLGAGAFKGVVEATAYGLIKSDAAM 616  
 QY 640 QVAVKMLKADSSEREALMSLKMWTQLGSHENIVNLLGACTLSGPIYLIFEYCCYGD 699  
 Db 617 TVAVKMLKPSAHLTREALMSLKLVSGLGNHNNIVNLLGACTVGGPTLVITEYCCYGD 676  
 QY 700 LNYLRKREKFRHTWTEIFKEHNFYPTFQSHPNSSMPGSRV-VQIHDPDSQISGLHGN 758  
 Db 677 LNFLRRKRSDF--IFSKEEQADAALYKNLLHSKSSCDSSNEYMDKPGVSVVYPTKTD 734  
 QY 759 SFHS---EDEIEYENOKRLEEDENLVTFEDLLCFAYQVAKGMFELEKFSVHRDLAAR 815  
 Db 735 KRRSARIDSIEDYERDVTPTAIMEDEL-ALDLELLSFSYQVAKGMFLASKNCIHRDLAAR 793  
 QY 816 NVLYTHGVKVKICDGLARDIMSDSNVYVGNARLPVKWMAPESEFEGYTIKSDVWSYG 875  
 Db 794 NILLHFRITKICDGLARDIRNDSNVYVGNARLPVKWMAPESEFNCYVTFESDVMWSYG 853  
 QY 876 ILLWEIFSLGVNPGIPVDANFYKLIQNGFKMDOPFYATEIYIIMQSCWAFDSRKPS 935  
 Db 854 IFLWELFSLGSSPYGMPVDSFYKMKIEGFRMLSPHEAPAMYEVMKTCWDADPLKRT 913  
 QY 936 FPNLTSFLGCOLADAEEAMYNQV 958  
 Db 914 FKQVVLIEKQISDSSKHIYSNL 936

RESULT 3  
 Q63116 PRELIMINARY; PRT: 978 AA.  
 AC Q63116;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE C-kit receptor tyrosine kinase.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY;  
 RX MEDLINE=92003944; PubMed=1912577;  
 RA Tsujimura T., Hirota S., Nomura S., Niwa Y., Yamazaki M., Tono T.,  
 RA Moril E., Kim H., Kondo K., Nishimune Y., Kitamura Y.;  
 RT "Characterization of Ws mutant allele of rats: A 12-base deletion in  
 RL Blood 78:1942-1946(1991).  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 DR EMBL: D12524; BAA02094.1; -  
 DR HSP; P11362; IFCG.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR003598; Ig\_c2.  
 DR InterPro: IPR003600; Ig-like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR001824; RtkkinaseIII.  
 DR Pfam: PF00047; Tvr\_pkinase.  
 DR Pfam: PF00069; pkinase; 1.

ProDom: PD000001; Euk\_pkinase; 2.  
 DR SMART: SM00408; Igc2; 1.  
 DR SMART: SM00410; Igc-like; 3.  
 DR SMART: SM00219; Tyrc; 1.  
 DR PROSITE: PS0107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE: PS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
 KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
 KW Phosphorylation; Receptor; Transmembrane;  
 KW Tyrosine-protein kinase.  
 SQ SEQUENCE 978 AA; 109341 MW; 0958C33F19889051 CRC64;

Query Match 24.0%; Score 1263.5; DB 11; Length 978;  
 Best Local Similarity 33.8%; Pred. No. 1.6e-95;  
 Matches 334; Conservative 171; Mismatches 361; Indels 121; Gaps 37;

QY 35 CVLINHKNDDSSVGKSSSYPMVSESPEDLGCALRPOSSGTVEAAAEVVDVSASITQVL 94  
 Db 12 CVLLVLLRGOTGTSQSPASGEPSP-----STPAQS-----ELIVEAGDTIRLT 57  
 QY 95 VDAIGNISLWFKHSLNQCQPHFDLQNRGVSVMLKMTF-TOAGEYLLFIQSEATNYT\_153  
 Db 58 CTDPAPVK--WTFE-----ILDVRIENKQSEWIREKAEATHGKYTCVSGSLRSSI 107  
 QY 154 ILFTVSIRN-TLLYTLRRPYFRKMNQDALVCISESVPEPIVWVLCDSOGESCKESPA 212  
 Db 108 YVF---VRDPAVLFVLGVLPLFGK-EDNDALVRCLPTDPQ-VSNYSLIECDGKSLPTDLKF 162  
 QY 213 V-----VKKEKVLHFLGTDIRCCARNELGRECTRLFTID---LNQTPOTTLPOL 260  
 Db 163 VNPKAGITIKNVKRAYHRLC---IRCAQREGKWHRSDFTLKVRRAIKAIPIVSVPE 219  
 QY 261 --FLKYGEPLWIRCKAVHNVHGFGLTW-----ELENKALEEGNYFEMS--TYSTNRTMIR 311  
 Db 220 SHLLKEGDTFTVICTIKDVSTSDSMWIKLNPOQSKAQVKRNSWHQGDENYERQELT- 278  
 QY 312 ILFAFVSSVARNDTGYTCSKSKHPQSALVT---IVGKGFN---ATNSEDYEIDQYE 365  
 Db 279 -----ISSARVNDSGVFMCYANNFTGSANVTTLKVVEKGFNIIPVKNNT--VFVTDGE 331  
 QY 366 EFCFSYFRKAYPO-IRCTWTFSRKSPCEOKGLD---NGYSISKFCNHHK-----OP 413  
 Db 332 NVDLVVEFEAYPKPEHQWIYMNRT--PTNRGDEYKSDNQSNIRYVNELRRLRLKTEG 389  
 QY 414 GEYIFHAENDDAOFTKMTLNIRRKQVLA--BASASQACSFSDGYPLPSWTWKKCDKS 471  
 Db 390 GTYTFVLSNDSVSASVTEDVYVNTKPEILTYDRLMNGRLQCVAAAGPPEPTIDWYFCTGAE 449  
 QY 472 PNCETEITE-GWNNRANKRVKGVQWVSSSTLNKSEAIKGLFLVKCCAYNSLGTSCETILLN 530  
 Db 450 QRCVTVPVPPVDQIQNASVSPGKLVQSSIDSSVFRHNGTVECKASNAVGS--SAFFN 507  
 QY 531 SPGPFPF-----IQDN-----ISFYATIGVCLLFIVLTLTLLCHIKYKKQFRYESOL 576  
 Db 508 ---FAKKNRKEIQPHLFTPTLLIGFVVVTAG--LMGIIVMWL--AYKYLOKPMYEVQW 559  
 QY 577 QWY-QVTGSSDNEYFYVDREYEDYDLKWEPPRENLEFGKVLGSGAFKGVNNAATAYGISKT 635  
 Db 560 KVVVEING---NNYVYIDPTQLPYDTHKWEPPRNRLSFGKTLGAGAFKGVVEATAYGLIKS 616  
 QY 636 GVSQVAVKMLKADSSEREALMSLKMWTQLGSHENIVNLLGACTLSGPIYLIFEYCC 695  
 Db 617 DAAMTVAVKMLKPSAHLTREALMSLKLVSGLGNHNNIVNLLGACTVGGPTLVITEYCC 676  
 QY 696 YGDLLNLYRSKREKFRHTWTEIFKEHNFYPTFQSHPNSSMPGSRV-VQIHDPDSQISG 754  
 Db 677 YGDLLNLYRSKREKFRHTWTEIFKEHNFYPTFQSHPNSSMPGSRV-VQIHDPDSQISG 754  
 QY 755 LHGNSFHS---EDEIEYENOKRLEEDENLVTFEDLLCFAYQVAKGMFELEKFSVHRD 811  
 Db 735 TKYDKRRSARIDSIEDYERDVTPTAIMEDEL-ALDLELLSFSYQVAKGMFLASKNCIHRD 793

```
QY 812 LAARNVLTHGRVVKICDFGLARDIMSDSNVYVGNARLPVKWMAPESELFEGYITIKSDV 871
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
794 LAARNLLTHGRITKICDFGLARDIRNDSNVYVGNARLPVKWMAPESEIFNCVYTESDV 853
QY 872 WSYGILLWEIFSLGNPNYPGIPVDANFYKLQNGFKMDOPFYATEIYIIMSCWAFDSR 931
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 854 WSYGLFELFELFSLGSSPYGPMVDTSKFKYMKIEGFRMLSPHAPAAVYEMKTCWDADPL 913
QY 932 KRPSPNLTSLFGCOLADAEAEAMQNV 958
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 914 KRPTKQVVOLIEKQISDSKHIYSNL 940
RESULT 4
ID O97744 PRELIMINARY; PRT; 964 AA.
AC O97744;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Mast/stem cell growth factor receptor (Fragment).
GN KIR1*0101.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAMPSHIRE BREED; TISSUE=LEUKOCYTE;
RX MEDLINE=98391767; PubMed=9724328;
RA Marklund S., Kiljas J., Rodriguez-Martinez H., Roennstrand L., Puna K.,
RA Moller M., Efors-Lilja I., Andersson L.;
RT "Molecular evolution of the dominant white phenotype in the domestic
RT pig.";
RL Genome Res. 8:826-833(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HAMPSHIRE BREED; TISSUE=LEUKOCYTE;
RA Marklund S.;
RL Thesis (1997), Department of Animal Breeding and Genetics,
RL Swedish University of Agricultural Sciences, Uppsala, SWEDEN.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CSP-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL: AJ223228; CAA11196.1; -
DR HSSP; P11362; IFGK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001824; RYKinaseIII.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 2.
DR SMART; SM00409; Ig_1; 1.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00219; TyrcK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00240; RECEPTOR_TYR_KIN_III; 1.
KW Glycoprotein; Immunoglobulin domain; Phosphorylation; Receptor;
KW Transmembrane; Tyrosine-protein kinase.
FT NON_TER 964
SQ SEQUENCE 964 AA; 108287 MW; 90E72EDFAB1358B7 CRC64;
```

```
Query Match 23.9%; Score 1262.5; DB 6; Length 964;
Best Local Similarity 33.4%; Pred. No. 1.8e-95;
Matches 332; Conservative 169; Mismatches 349; Indels 143; Gaps 38;
QY 47 VGGSSSPWVSESPDLGCLALPQSSGTVYEAARVVDVYSASITLQLVLDAPGNISCLMW 106
```

```
RESULT 5
Q9TQ01
ID Q9TQ01 PRELIMINARY; PRT; 964 AA.
AC Q9TQ01;
DT 01-MAY-2000 (Tremblrel. 13, Created)
```

```
DB 20 VQTGSSQPSV--SPEEL-----SPPSIHAKS--ELIVSAGDEIRLFCDPGSVK--WT 67
QY 107 FRHSSLNCOPHEDLQNRGVVSWILKMTETQAGEYLLFIQSEATNYTILFTVSIRN--TLL 165
DB 68 FETLG-----QLSENTHAETWIEKAEAMNTGNYTCTNEGGSSSYVF---VRDPEKL 117
QY 166 YTLRRRPFYFRKMEQDALVCTISVPEPIVWELCDOSGESCKBESPAV-----VKKE 217
DB 118 FLVDDPLYGK-EDNDALVRCPLDPE-VTNYSLTGEGKPLPKDLTFVADPKAGITIRNV 175
QY 218 EKVLFELFCTDIRC--CARNELGRE-CTRLFTID----LNQTPOTTLPO--LFLKVGSEPL 268
DB 176 KREYHRL-----CLHCSANQGGKSVLSKFTLKVRAAIRAVPVVAVSKASYLLREGEF 229
QY 269 WIRCKAVHVNHGFLTWELN---KALEEGNYFEMSTYSTNRTMIRLLPAFVSSVARNDT 325
DB 230 AVNCLIKDVSSSVSDSWIRENSQTKAQVKNRNSWHQGFN----FLROERLTISSARVND 285
QY 326 GYVTCSSSKHPQSALVT---IVGKGFINA-----TNSSEYVEIDQYEEFCFSVRF 373
DB 286 GYFMCYANNTFGSANVTTTLEVVDKGFINIFPMNNTTVFVNDGEDVDL-----IVEY 337
QY 374 KAYPQ-----IRCTWTFSRKSPFCEQKGLDNGYSISKFCNHKH-----QPGEYIF 418
DB 338 EAYPKPEHRQWIYMNRTATDKWEDYPKSE-----NESNIRYVSELHLTRLKGTGGTYTF 392
QY 419 HAENDDAQTKMETLNIIRKPKQVLA--EASASQASCFSDGYPLPSWTWKKCSKSPNCTE 476
DB 393 LVSNDVNSVSTENVYVNTKPEILTHDRLMNGMLQCVAAGPEPTIDWYFCPGTEQRCVS 452
QY 477 EI-TEGVNRKANRKVEGOWSVSSSTLNMSPAIKGFLVKCCAYNSLGTSCETILLNSPGPF 535
DB 453 PGPVVDVQIONSSVPFGKLVIIHSSIDYSAFKNGTVCEARNDVGVKS---SAFNF----F 506
QY 536 PFIQD-----NISFYATIGVCLLFTIVLTLTLLCHRYKKQFYESOLOMW-QVTGS 584
DB 507 APKEQIHAHTLFTPLLIGFVIAAGMCMIIWMILT---YKLOKPMYEVQWKVVEEING- 561
QY 585 SNEYFYVDFREYEDLKWEFFPRENLEFGKVLGSGAFGKVMNATAYGISTGVSIOVAVK 644
DB 562 --NNYVYIDPTQLPYDHKWEFFPRNRLSFGTKLGAGAFGKYVEATAYGLIKSDAAMTAVK 619
QY 645 MLKEKADSSEREALMSELMKMTGLSHENIVNLLGACTLSGPIYLLIPEYCCYGDLLNLYR 704
DB 620 MLKPSAHLTEREALMSELKVLGYLGNHMINVNLGACTIGGPTLVITEYCCYGDLLNLYR 679
QY 705 SKREKEFRTWTE-----IFKEHNFSPYPTFQSHPNSSM---PGSREVQIHPDSDQISGL 755
DB 680 RKRDSFICSKQEDHAEALYKLLHKSKESSCSDSTNEYMDMKPGVSYV-VPTKADKRSA 738
QY 756 HGNSFHSDEIEIYENOKRLEEEEDLNVLTFEDLLCFAYQVAKMEFLFKSCVHRDLAAR 815
DB 739 RIGSY-----IERDVTPTAIMEDEL-ALDLEDLSFSYQVAKGMAFLASKNCIHRDLAAR 792
QY 816 NVLVTHGKVVKICDFGLARDIMSDSNVYVGNARLPVKWMAPESELFEGYITIKSDVHSYG 875
DB 793 NILLTGHRITKICDFGLARDIKNDSNVYVGNARLPVKWMAPESEIFNCVYTESDVHSYG 852
QY 876 ILLWEIFSLGNPNYPGIPVDANFYKLQNGFKMDQPFYATEIYIIMSCWAFDSRKRPS 935
DB 853 IFLWELFSLGSSPYGPMVDTSKFKYMKIEGFRMLSPHAPAAVYEMKTCWDADPLKRPT 912
QY 936 FPNLTSLFGCOLADAEAEAMQNVQVGRVSEC-PH 967
DB 913 FKQIVOLIEKQISESTNHIYSN----LANCSPH 941
```



		Query Match	23.9%; Score 1261.5; DB 6;	Length 964;
		Best Local Similarity	33.4%;	Pred. No. 2.2e-95;
	Matches	332; Conservative	169; Mismatches	349; Indels 143; Gaps
QY	47	VGRSSSYPMVSESPEDLCALRPOSGTGVYEAAAVEVDVSASITLQVLVDAPGNISCLWV	106	
Dd	20	VQTGSSQPSV--SPEEL-----SPPSIHPAKS-ELIVSAGDEIRLFCTDPGSVK--WT	67	
QY	107	FKHSSLNCQPFDQLONRGVSMWLKMTETQAGEVLLFIQS EATNYTTLTFTVSRN-TLL	165	
Dd	68	FETLG-----QLSENTHAEIWEKEAMTMNGNTCTNEGGLSSIYP---VRDPEKL	117	
QY	166	YTLRPRPYFRKMENODALVCITESYPEPIVEVWLDCDSOGESCKESPAY-----VKRE	217	
Dd	118	FLVDDPLYLK-EDNDALVRCP-LTDPE-VTNYSLTCCCEGPKPLDKLTFVADPKAGITIRNV	175	
QY	218	EKVLUHELFGTDIRC-CARNELGRE-CTRLFTID----LNQTPOTTL PQ--LFLKVGEP L	268	
Dd	176	KREYHRL-----CLHC SANQGSKSVLSKKF TL KLVRAAIRAVPVVA VASKASY LLREGEE F	229	
QY	269	WI RCKAVHVNHGFGLTWLEIN---KALEEG NYFMESTYSTNR T W I RLFA PV SSVAR NDT	325	
Dd	230	AVMLIKDVSDSSVMIRENSQPAOKVRNSWHGGDFN----FLROERLARISSA RV DS	285	

RA Marklund S., Kijas J., Rodriguez-Martinez H., Roennstrand L., Funa K.,  
RA Moller W., Edfors-Lilja I., Andersson L.;  
RT "Molecular evolution of the dominant white phenotype in the domestic  
pig.";  
RL Genome Res. 8:826-833(1998).  
CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
CC -I- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-  
CC PROTEIN KINASES.  
DR EMBL; AJ223230; CAAL1198.1; -.  
DR HSSP; P11362; IFKG.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR003599; Ig\_Like.  
DR InterPro; IPR003600; Ig\_Like.  
DR InterPro; IPR003606; Ig\_MHC.  
DR InterPro; IPR001824; RTKinaseIII.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00047; Ig\_1.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 2.  
DR SMART; SM00409; IG; 1.  
DR SMART; SM00410; IG\_Like; 1.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
KW Glycoprotein; Immunoglobulin domain; Phosphorylation; Receptor;  
KW Transmembrane; Tyrosine-protein kinase.  
FT NON\_TER 964 964  
SQ SEQUENCE 964 AA; 108287 MW; 1B21292A962E9191 CRC64;

Query Match 23.9%; Score 1261.5; DB 6; Length 964;  
Best Local Similarity 33.4%; Pred. No. 2.2e-95;  
Matches 332; Conservative 169; Mismatches 349; Indels 143; Gaps 38;

QY 47 VKSSYPVMSVSPEDLCALRQSSGTVEAAAEVDSASITLQVLDAPGNISCLWV 106  
DB 20 VQTGSQSPV--SPEL-----SPSTHPAKS-ELIVSAGDEIRLFCFDPGSVK--WT 67  
QY 107 FKHSLENCQPHFQDQNRGVVSMVILKMTETQAGEYLLFIQSEATNTYLTFTVSRN--TLL 165  
DB 68 FETLG-----QLSENTAEWIEKAEAMNTGNYCTNEGLSSSIYVF---VRPEKL 117  
QY 166 YTLRRPYFRMENODALVCISESPPIVWVLCDSQGESCKEESPAV-----VKKE 217  
DB 118 FLVDPPLYCK-EDNDALVRCLTDPDE-VTNYSLTGCEGKPLPKDLTFVADPKAGITIRNV 175  
QY 218 EKVLEHFGTDRC--CARNELGRE-CTRLFTID-----LNQTFQTLPO--LFLKYGEPL 268  
DB 176 KREYHRL-----CLHCSANQGGKSVLSKFTTLKVRRAIRAVPVAVSKASYLLREGEEF 229  
QY 269 WIRCKAVHNVHGFGLWLEEN-----KALEENGYEMSTYSTNRMTIRILFAFVSVAARNDT 325  
DB 230 AVMLIKDVSSSVSDSMWIRENSTQAKVKNSSHQGFN----FLROEKLITISSARVND 285  
QY 326 GYYTCSSKHPSQALVT---IVKGKFINA-----TNSSEYDIDYEEFCFSVRF 373  
DB 286 GVEMCYANNTFGSANVTTLLEVVDKGFNIFPMWNTTVFVNDGEDVDL-----IVEY 337  
QY 374 KAYPQ-----IRCTWTFSRKSFCEQKGLDNGYSISKFNHK-----QPGYIIF 418  
DB 338 EAYPKPEHQWIMYMNRTATDKWEDPKSE-----NESNIRYSELHLRLKLGTEGGYTF 392  
QY 419 HAENDDAQFTKMTFLNIRKQVOLA--EASASOASCFSDGYPPLPSWFWKCKSDKSPNCTE 476  
DB 393 LYSNADVNSSTVYVNTKPEILLTHDLRMNGMLQCVAAAGFPEPTIDWYFCPTGEQRCV 452  
QY 477 EI-TEGVWNRKANRKYFGQWSSSTLMSPAIKGFLVKCCAYNSLGTSCFTILLNSPGPF 535  
DB 453 PVGPVDVQIONSSVSPFGKLVIHSSIDYSAFKHNGTVECRAYNDVGKS--SAFEN----F 506  
QY 536 PFIQD-----NISFYATTGCVLLFTVLTLTLLCHYKKQFYRESQLQWV-QVTGS 584

DB 507 AFKEQTHAHTLTPPLIGFVIAAGMCIIVMILT-----YKLOKPMYEVQWVVEELNG- 561  
QY 585 SDNEYFYVDFREYEDLKWEFFPRENLEFGLGSGAPGKVMNATAYGISKTGVSQVAVK 644  
DB 562 --NNYVVIDPTQLPYDKHWEFFPRNLSFGKTLGAGAFGKVVEATAYGLIKSDAAMTVAVK 619  
QY 645 MLKEKADSSEREALMSKMTQLGSHENIVNLLGACTLSGPIYLIEFYCCYGDLLAYLR 704  
DB 620 MLKPSAHLTEREALMSKLVLSYLGHNHMTNVLGACTIGGPTLVITVEYCCYGDLLNPLR 679  
QY 705 SKREKFRHTWTE-----IFEKHNFSFYPTFQSHPNSSM---PGSREVOIHPDSDQISGL 755  
DB 680 RKRDSTFCSKQEDHAEALYKLLHLSKESCDSTNEYMDMKPGVSYV-VPTKADKRRA 738  
QY 756 HGNSFSEDEIEYENQKRLEEDLANVLTFFEDLLCFAYQVAKGMFELEFKSCVHRDLAAR 815  
DB 739 RIGSY-----IERDVTPAIMEDDEL-ALDLEDDLFSYQVAKGMAFLASKNCIHRDLAAR 792  
QY 816 NVLVTGKVVKICDFGLARDIMSDSNVYVGRNARLPVKNWAPESLFEPIYTIKSDVWSYG 875  
DB 793 NILTHGRITKICDFGLARDIKNDSNVVGRNARLPVKNWAPESIFNCVYTFESDVWSYG 852  
QY 876 ILLWEIFSLGVNVPVGPVDFANFYKLIQNGFKMDOPFYATVEEYIIMOSCWAFDSRKRPS 935  
DB 853 IFLWELFSLGSSPYGPMFVDSKFKYKMIKEGFRMLSPHAPVEMYDIMKTCWDADPLKRPT 912  
QY 936 FPNLTSLFGQLADAEAEAMYQNVDGRVSEC-PH 967  
DB 913 FKOIVQLIEKQISESTNHIIYSN----LANCSPH 941  
RESULT 7  
Q8WN23 PRELIMINARY; PRT; 979 AA.  
ID Q8WN23;  
AC Q8WN23;  
DT 01-MAR-2002 (T-EMBLrel. 20, Created)  
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
DE C-KIT.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN .[1]  
RP SEQUENCE FROM N.A.  
RA Zemke D., Yuzbasivan-Gurkan V.;  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF448148; AAL40833.1; -.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR003600; Ig\_Like.  
DR InterPro; IPR003606; Ig\_MHC.  
DR InterPro; IPR001824; RTKinaseIII.  
DR InterPro; IPR002290; Ser\_thr\_pkinase.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00047; Ig; 2.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 2.  
DR SMART; SM00409; IG; 4.  
DR SMART; SM00410; IG\_Like; 2.  
DR SMART; SM00220; S\_TKc; 1.  
DR SMART; SM00219; Tyrc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; UNKNOWN\_1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; UNKNOWN\_1.  
DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; UNKNOWN\_1.  
SQ SEQUENCE 979 AA; 109753 MW; 46C30D5DEB8E33D3 CRC64;

Query Match 23.7%; Score 1250.5; DB 6; Length 979;  
Best Local Similarity 33.6%; Pred. No. 1.9e-94;  
Matches 334; Conservative 161; Mismatches 345; Indels 153; Gaps 39;



```
Db 278 SSARVNDGVMFCYANNFTGSAVNTTLEVVVKGFINIFPMINTVFVNDGENVDL---- 333
Qy 366 EFCFSVRKAYPO-IRCTWTFSRKSP--PCEQKGLDNGYSISKFCNHK-----QPGE 415
Db 334 ----IVEYEAFPEHQOVIYMRFTDKWEDYPKSENIYRVSELHLRLKLGTEGT 389
Qy 416 YIFHAENDDAQFTKMTLNIRRRPOVLA--EASASQACSFSDGYPLPSWTWKCKSDKSPN 473
Db 390 YTLVNSDVNAAIAFNYYVNTPEILTYDRLVNGMLQCVAAGPEPTIDWFCPGTEQR 449
Qy 474 CTEEITE-GVNNRKANKRVFGQVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLNSP 532
Db 450 CSASVLPVDVOTLNSGPPGKLVQSSIDSSAFKNGHTVECRAYNDVGT--SAYEN-- 505
Qy 533 GPPFPFTQD-----NISFVATICVCLLFIVVLLTLLCHKKYKQFVRESQOMV-QV 581
Db 506 --FAFKEQIHPHTLFTPLLLGFGVIVAGMCIIVMLT-----YKYLQKPMVEQVKKVEEI 559
Qy 582 TGSDDNEYFVDFREYEDLKWEFPRENLEFGKVLGSGAFKVMNATAYGISKTGVSIQV 641
Db 560 NG---NNYVVIDPTQDPYDHKWEFPNRNLSEFGTLGAGATKVVVEATAYGLIKSDAAMTV 616
Qy 642 AVKMLKEKADSSREALMSLKMTOQLGSHENIVNLLGACTLSGPIYLIFEYCYGDLN 701
Db 617 AVKMLKPSAHLTEREALMSLKVSLYLGNNHINIVNLLGACTIGGPTLVITEYCCYGDLN 676
Qy 702 YLRKREKFRHTWTE-----IKENHFSYPTQSHPNSSM----PGSREVQIHPDSQI 752
Db 677 FLRRKRDSFCSQEDHAEALYKNNLLHSKSSDSTNEYMDMKPGVSYV-VPTKADK 735
Qy 753 SGLHGSFHSDETEYENOKRLEBEEDLNVLTPEDLLCFAYOVAKGMFELEKFCVHRDL 812
Db 736 RSVRIGSY-----IERDVTPEIMDEDEL-ALDLEDLLSFSYQVAKGMAFLASKNCIHRDL 789
Qy 813 AARNVLVTHGVKVKICDFGLARDTMSDSNVYVRGNARLPVKWMAPESELFEIYTIKSDVW 872
Db 790 AARNILLTHGRITKICDFGLARDIKNSNVYKGNARLPVKWMAPESEFNCVYTFESDVW 849
Qy 873 SYGILLWEIFSLGVNPNYPGIPVDANFYKLQNGKPMQOPFYATEEYIIMOSWAFDSRK 932
Db 850 SYGIFLWELFSLGSPYPGMPVDSKFKYMIKEGFRMLSPHAPAEYDIMKTCWDADPLK 909
Qy 933 RPSFPNLTSLGCOLADAEAEAMYNQV 958
Db 910 RPTFKQIVOLIEKOISESTNHHYSNL 935

RESULT 9
Q9XS93 PRELIMINARY; PRT: 978 AA.
AC Q9XS93;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE K.T.
GN C-KIT.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CEREBELLUM;
RA London C.A.; Galli S.J.; Yuuki T.; Hu Z.-Q.; Helfand S.C.;
RA Geissler E.N.;
RT "Spontaneous canine mast cell tumors express tandem duplications in
RT the proto-oncogene c-kit.";
RL Exp. Hematol. 0:0-0(1999).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL; AF099030; AAD28369.1; -.
DR HSP; P11362; 1FGK.
```

```
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003599; Iq.
DR InterPro: IPR003600; Iq_like.
DR InterPro: IPR003006; Iq_MHC.
DR InterPro: IPR001824; RtkKinaseIII.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; Iq; 2.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 2.
DR SMART: SM00409; IG; 2.
DR SMART: SM00410; IG_Like; 1.
DR SMART: SM00219; TyKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Phosphorylation;
KW Receptor; Transferase; Transmembrane; Tyrosine-protein kinase.
SQ SEQUENCE 978 AA; 109651 MW; 5AC31E5AC4E9910F CRC64;

Query Match 23.6%; Score 1244; DB 6; Length 978;
Best Local Similarity 33.2%; Pred. No. 6.4e-94;
Matches 330; Conservative 161; Mismatches 346; Indels 158; Gaps 38;

Qy 47 VGKSSYPMVSESPEDLGCALRPOSSGTVYEAAAVEVDVSASITLQVLVDAPGNISCLMW 106
Db 22 VQTGSSQPSVSPGPELP-SIHPAKS-----ELIVSVGDELRLSLCTDGFVK--WT 69
Qy 107 FKX-SSLNCOPHDLQNRGVVSMVLKMTQOAGBYLLFIOSEATNYTIL-----FT 157
Db 70 FETLQLGNETH---NEWITE-----KAEAG-----HTGNVTCNTRDGLRSIV 110
Qy 158 VSIRN-TLLYTLRRPYFKMENQDALVCISEVPPIVWVLCDSQGSCKEESPAV--- 213
Db 111 VFVRDPAKLFVLVDLPYLGK-ECNDTLVRCPLTDP-VTNYSLRGCEKPLPKDLTFVADP 168
Qy 214 ----VKEEKVLHELPGTDIRC--CARNELGRE-CTRLEFTID----LNQPTTLPOL- 260
Db 169 KAGITIRNVKREYHRL-----CLHCSADQKGRVLSKFKTLKVRAAIRAVPVSVSKTS 222
Qy 261 -FLKVGPELWIRCKAVHNVHGFGLTWELN---KALEEGNYFEMSTYSTNTRMIRLFAF 316
Db 223 SLLKEGEAFSVMCFIKOVSEFVDSMWIKENSQTNATQSNHMHGDFNFERQEKLI---- 278
Qy 317 VSSVARNDTGYTCTSSSKHPQSALVT---IVGKFINA-----TNSDEYIDQY 364
Db 279 ISSARVNDGVMFCYANNFTGSAVNTTLEVVVKGFINIFPMINTVFVNDGENVDL---- 335
Qy 365 EFCFSVRKAYPO-IRCTWTFSRKSPCEOKGL---DNGYSISKFCNHK-----QP 413
Db 336 ----IVEYEAFPEHQOVIYMRFTDKWEDYPKSDNESNI-RYVSELHLRLKNGEG 389
Qy 414 GEYIFHAENDDAQFTKMTLNIRRRPOVLAESAASOA--SCFSQGYPLPSWTWKCKSDKS 471
Db 390 GTYTFQVNSDVNSSVTFNVYVNTKPEILLTHESLTNGMLQCVAAGPEPAVDWFCPGA 449
Qy 472 PNCTEEI-TEGVNWRKANRVFGQVSSSTLNMSAIGFLVKCCAYNSLGTSCETILLN 530
Db 450 QRCSPVIGPMQVQMONSSLSPSGKLVQSSIDYSFAKNGTVECRAYNNVGRS--SAFE 507
Qy 531 SPGPPFPFTQD-----ISFVATICVCLLFIVVLLTLLCHKKYKQFVRESQ 576
Db 508 ----FAFKGNSKEQIHPHTLFTPLLLGFGVIVAGMCIIVMLT-----YKYLQKPMVEQV 559
Qy 577 QMV-QVTGSSDNEYFYDFREYEDLKWEFPRENLEFGKVLGSGAFKVMNATAYGISKT 635
Db 560 KVVEEBING---NNYVVIDPTQDPYDHKWEFPNRNLSEFGTLGAGATKVVVEATAYGLIKS 616
Qy 636 GVSIOVAVKMLKEKADSSREALMSLKMTOQLGSHENIVNLLGACTLSGPIYLIFEYCC 695
Db 617 DAAMTVAVKMLKPSAHLTEREALMSLKVSLYLGNNHINIVNLLGACTIGGPTLVITEYCC 676
Qy 696 YGDLNLNLRSKREKFRHTWTEIFKEH-NFSFYPTFQSHPNSSMPGSRREVQIHPDSQISG 754
```



```

RESULT 11
O77589 PRELIMINARY; PRT; 945 AA;
AC O77589;
DT 01-NOV-1998 (TREMELrel. 08, Created)
DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
DE Tyrosine kinase receptor homolog (Fragment).
GN KIT.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Lear T.L., Adams M.H., McDowell K.J., Ponce de Leon F.A.,
RA Sullivan N.D., Coogle L., Ferguson E., Chambers T.M., Bailey E.;
RT "Chromosomal location of genes for ESR, ETS2, KIT, MX1 and PGR
RT in the horse, Equus caballus.";
RL Proc. Xth Amer. Coll. Dom. Anim. Cytogenet. Gene Mapp. 0:0-0(1997).
RN [2];
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Rieder S., Checa-Cortes M.L., Joerg H., Stranzinger G.;
RT "An Equine sequence homologous to the c-KIT proto-oncogene mapped to
RT chromosome 3q13-21.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
DR EMBL: AF055037; RAC23611.1; -.
DR HSSP: P11362; IFGR.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR001824; RTKinaseII.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00047; Ig. 1.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Euk_pkinase; 2.
DR SMART: SM00409; Ig. 2.
DR SMART: SM00219; TyrcKc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
KW Glycoprotein; Immunoglobulin domain; Kinase; Phosphorylation;
KW Receptor; Transmembrane; Tyrosine-protein kinase.
FT NON_TER 945
SQ SEQUENCE 945 AA; 106143 MW; 7DB424A89DF5C340 CRC64;

Query Match 23.3%; Score 1229; DB 6; Length 945;
Best Local Similarity 32.5%; Pred. No. 1 le-92;
Matches 323; Conservative 166; Mismatches 368; Indels 136; Gaps 31;

QY 47 VKKSSYPWSEPEDGALRPSQSGTYVEAAAEVDVSASITLQVLVDAPGNISCLWV 106
DB 20 VQTGSSQPSV--SPGEL-----SPPSIHPAKS--ELIVSGVEIRLLCADPGFVK--WT 67
QY 107 FK-HSSLNCPHFQDLQNRGVSVILKMTETQAGEVLLFIQSEATNYTILFTVSIINTLL 165
DB 68 FETWGLSENTHKE-----WVTEKAEATNTGVTCTNEGGLSSIVVF---VRDPAK 116
QY 166 YTLRRYFRKMNODALVCISESPPIVVEWVLCDSOGESCKEESPV-----VKKE 217
DB 117 LFLFDPSLYKGESSDVLRCPLTDPE-VTNYSLMACEGKSLPKDLTFVADPKAGITRNW 175
QY 218 EKVHLEFGTDIRCCARNELGRECTRLFTID----LNQTPQTTLPO--LFLKVGEPWLIR 271
DB 176 KREYHRLC---LRCSADKDGKSVLSNFTLKVRAAIRAVPVVSVKASYLLREGEEFSVT 232

```

```

QY 272 CKAVHVNHGFLTWELNKALEEGNYFEMSTYSTNRTMIRLPAFVSSVARNDTGYYTCS 331
DB 233 CLIKDVSSVDSVDSWIRENSRTKE-QVKSSSSHQDGFNFVQRERLTISPARNVDSGVFWCY 291
QY 332 SSKHPSOSALVT---IVKGGFINA-----TNSSEDEIQQVEFCFSVRFKAYPQ- 378
DB 292 ANNTFGSANVTTLLEVLDKGCIINVFPMMNTTVFVNDGENVDL-----IVEYESYKPK 343
QY 379 -----IRCTWTFSRKSPFCEQKGLDNGYISIKFCNKKH-----QPGEYIIFHAEND 424
DB 344 EHOQWIMYNTSTDKWEDYPKSE-----NESNIRYSELHLRLKLGTEGGTYTLFVNSND 398
QY 425 AQTFTKFTLNIIRKKPOVLA--EASASQASCFSDGYPPLPSMTWKCKSDKSPNCTEEI-TEG 481
DB 399 VDSSTVFNVYNTKPEILTRDLRMNGMLQCVAAAGFPPEPTIDWYFCPTEQRCSPVGPVD 458
QY 482 VWRKANRKYFGQWVSSTLNMSEAIGKFLVKCAYNLSGTSCETILLNSGPPFPFIDON 541
DB 459 VKIQNSSVSPFGKLVQSSIDYSAFKHNGTVCERAYNDVGKSSAFF-----N 505
QY 542 ISFYATIGVCLLF-----IVVLTLLCHYKKQFRYESOLOMW-QVTGSSDN 587
DB 506 FAFKEQIHPTLFTPLLLIGSVAAAGMCMVIMVLTLYKLOKPMYEVOMKVVEEING---N 562
QY 588 EYFVVDPREYEDLKWEFPRENLEFGKVGAFKVMNATAYGISKTGVSIOVAVKMLK 647
DB 563 NYVIIDTQPLQDHKWEFPNRNLSFGKTLGAGAFKVVETATAYGLIKSDAAMTVAVKMLK 622
QY 648 EKADSSERREALMSELKMTQLGSHENIVNLGACTLSGPIYVLIFFEYCYGDLNLYRSKR 707
DB 623 PSAHLTEREALMSELKVLVSLGNHMNIVNLLGACTVGGPTLVITEYCYCGDLLNLRKR 682
QY 708 EKPHRTWTE-----IPKEHNFSEYPTFQSHPNSSM---PGSREVQIHPSDDQISGLHN 758
DB 683 DSFTCSKQEDHAEAAALYKLLHSSKESCDNSTNEYMDMKPGVSYV-VPTKTKRRAARIG 741
QY 759 SFHSEDEIEYENOKRLEEDLNVLTFEDLLCPAYQVAKGMERLEEFKSCVHRDLAARNVL 818
DB 742 SY-----IERDVAPSIMEDDEL-ALDLEDLLSFQVAKGAFNASKNCIHRDLAARNIL 795
QY 819 VTHGKVKIKCDFGLARDIMSDSNVVRGNARLPVKWMAPESLFEGIYTIKSDVMSYGILL 878
DB 796 LTHGRITKICDFGLARDIKDNSYVVKGNVRLPVKWMAPESIFNCVYTFESDVMSYGIFL 855
QY 879 WEIFSLGVNYPGIPVDANFYKLIQNGFKMDQPPYATEEYIIMQSWAFDSRRKRPSPFN 938
DB 856 WELFSLGSSYPGMPVDSKFKYKIKEGRMLSPHAPAEYDINKTCWDADPLKRPTFKQ 915
QY 939 LTSFLGCOLADAEEMAYQNVDRGVSECPHYQON 971
DB 916 IVOLIEKQISDSTNQIYSN-----LANCSPROEN 944

RESULT 12
Q98SU3 PRELIMINARY; PRT; 974 AA.
AC Q98SU3;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMELrel. 21, Last annotation update)
DE Receptor tyrosine kinase Fms.
GN FMS.
OS Danio dangila.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=127599;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE=21359118; PubMed=11466528;
RA Parichy D.M., Johnson S.L.;
RT "zebrafish hybrids suggest genetic mechanisms for pigment pattern
RT diversification in Danio.";

```

Db	470	EGESFGSGVGVESVLTVPGN--RRWTVVAVFNVLVGCGSDTFSMD-----VSDQIFTSAM	521
Qy	548	IG--VCLLFVTLTLLCHIKYKQFRYESQLQWQVOTGSSDNEYEVDFREYEVYDKWEF	605
Db	522	CGSTVAMVVLGLLIPWIKYKQKPYEIRKLIET--NGNYYTFIDPTQLPYNEKWEF	579
Qy	606	PRENLEFGKVLGSGAFGKVMNATAYGISKTGYSIQVAVKMLKEKADSSEREALMSLKMM	665
Db	580	PRDKLKLGKTLGAGAFKGVVEATAYGLGKEDNITRVAVKMLKASAHDPDEREALMSLKIL	639
Qy	666	TOLGSHENIVNLGACLTSPGIVLIEFYCCYGDLLNLRSKREKHFHRTWTEIFKEHNESF	725
Db	640	SHLGQHKNIIVNLGACHTGGPVLIVITEFYCCHGDLNLFKSAENF-----LNF	687
Qy	726	YPTFQSHPSNSMP-----GSREVOIHPDSQIISGLHGNSFHSDEIYEYNQKRLE-	777
Db	688	VMTIPDPPEMTDYKNISTERMFEVSDSGISSAC-----SDHYLDMRPVTSRPTLDSSECQ	744
Qy	778	EDLNVLTFEDLLCFAYQVAKGMFEFLFKSCVHRDLAARNVLVTHGKVKICDFGLGIARDM	833
Db	745	EDSPLDMDDLLRFSFQVAGQLDFLAAKNCIHRDVAARNVLLTNSRVAKICDFGLGIARDIM	804
Qy	838	SDSNVYVGRNARLPVKWMAPESLFEGIYTIKSDVNSYGYILLWEIFSLGVNPPGIPVDAN	897
Db	805	NDSNVYVGRNARLPVKWMAPESEFECVYTVQSDWSYGYIMLWEIFSLGKSPYPNVLVDSK	864
Qy	898	FYKLIQNGFKMDQPFYATEBIYIIMQSCWAFDSRRKRPSPNLTSLFGCOLADAEAM---	954
Db	865	FYKIKGCGYQMSRDFAPPMTYIMKWCNWLDAERTPTFSKISQMIQMLGETPEQOETQ	924
Qy	955	-YQNV-----DGRVSECPHTYQNRPF-----SRMDLGLLSP	986
Db	925	EYKNIPSEAEATOLESDDPAKHEDSEFETSCDQEEEDQPLMKP	968
RESULT 13			
Qy1909	AD	Q91909	PRELIMINARY; PRT: 954 AA.
IC	ID	Q91909	
DT	DT	01-NOV-1996 (TrEMBLrel. 01, Created)	
DT	DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)	
DT	DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)	
DE	DE	C-kit-related kinase 1 (xkrk1) precursor.	
OS	OS	Xenopus laevis (African clawed frog).	
OC	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;	
OC	OC	Xenopodinae; Xenopus.	
OX	OX	NCBI_TaxID=8355;	
RN	RN	[1]	
RP	RP	SEQUENCE FROM N.A.	
RX	RX	MEDLINE-95344996; PubMed-7619732;	
RA	RA	Baker C.V., Sharpe C.R., Torpey N.P., Heasman J., Wylie C.C.;	
RT	RT	"A Xenopus c-kit-related receptor tyrosine kinase expressed in	
RT	RT	migrating stem cells of the lateral line system.";	
RL	RL	Cell. Dev. 50:217-228(1995).	
CC	CC	!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).	
CC	CC	!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE	
CC	CC	PROTEIN KINASES.	
DR	DR	EMBL; Z48770; CAA88688.1; -.	
DR	DR	HSSP; P11362; 1FGK.	
DR	DR	InterPro; IPR000719; Euk_pkinase.	
DR	DR	InterPro; IPR003600; Ig.Like.	
DR	DR	InterPro; IPR003006; Ig.WHC.	
DR	DR	InterPro; IPR001824; RTKinaseII.	
DR	DR	InterPro; IPR001245; Tyr_pkinase.	
DR	DR	Pfam; PF00047; Ig. 2.	
DR	DR	Pfam; PF00069; pkinase; 1.	
DR	DR	ProDom; PD000001; Euk_pkinase; 2.	
DR	DR	SMART; SM00410; IG_Like; 2.	
DR	DR	SMART; SM00219; TyrKc; 1.	
DR	DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.	
DR	DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.	
DR	DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.	

DR PROSITE; PS00240; RECEPTOR\_TYR\_KIN\_III; 1.  
KW ATP-binding; Glycoprotein; Immunoglobulin domain; Kinase;  
KW phosphorylation; Receptor; Signal; Transmembrane;  
KW Tyrosine-protein kinase.  
FT SIGNAL 954 AA POTENTIAL.  
SQ SEQUENCE 954 AA 106859 MW; 9748894CBE0B537D CRC64;

Query Match 23.2%; Score 1226; DB 13; Length 954;  
Best Local Similarity 33.1%; Pred. No. 1.9e-92;  
Matches 305; Conservative 162; Mismatches 322; Indels 132; Gaps 30;

QY	107	FKHSLNCQPHFDLQNRGVVS-----WVLKMTETQAGEYLLFIQSEATNYTLFTVSI	160
DB	52	FOKSLMKKPR-DLKRP LNNSSETDOFFWIIKADLRHIGRICTNFTQENTSV--SLFV	108
QY	161	RNTLLYTRRPFYRKEMENODALVCISSEPEPIVWVLCDOSGESCKE-----ESPA	212
DB	109	KDAPRPFDLTFFIDVTEGADTVGMCPTDPMDDIAIEKCD--GSPLEPNFTTTDIEAGI	166
QY	213	VVKKEKVLHELFGTDIRCCARNELGR-ECTRLTFTIDLNQTTQLPQLFL-----KV	264
DB	167	TIKTVOLAFDSCY-----VCSGKNKSTVKKSSTFSTHVRPVK-KVPTVFLSKSRQLVKT	220
QY	265	GEPLWIRCKRAVHNHVGFLTW-ELENKALBEGNYFEMSTYSYTRNRMIRILFAFVSSVARN	323
DB	221	GEBFEVTCALVDVSTVKAQWLDEGVTKQANFRSSNNFYNLTKS-----DCVPYS	274
QY	324	DTGYTYCSSKHPSQ---SALVTIVGKGFINAT-NSSEDEYEDQVEEFCFSVRKAYPQ-	378
DB	275	ESRTFQOENALGOVYNATFTLDVIGVYNLTVENTTISYNAGNLVLKYIIDAYPHP	334
QY	379	IRCTWTSRKSPFCEQKGLDNGYSIK-----FCNHKH-----QCEYIFHAENDDA	425
DB	335	DDGWYTFYNETL----LNTSDHVATKDEGNRRYVSELHLRLKTEKGVVTFYTTNSDD	390
QY	426	QFTKMTLAIIRKPKVLAERASQAS--CFSDGYPLPSWTWKKCSKDSKPNCTEETEGWV	483
DB	391	DASVSNIOVKRPELIIAERTSEGTLQCVATGFVPJAIQWTFCEPGBQRCDTDPPLSPV	450
QY	484	NRK--ANRKVFQWVSSTLNKSEAIKGLFKCCAYNSLGTSCETILLNSPPFPF-TQD	540
DB	451	NEKFIQENSGLRIVVESTIDVNDLKNKTQCVASNEVESAYS-----FSPAKE	502
QY	541	N-----ISFVATTGCVLLFTVYVLTLLIHCYKKQFYESQLQWV-QVTGSSDNE	588
DB	503	KLKTHLTPLLIGFTAAAGLMCIAVAVL---MYKLOKPKYEIQWKVVEEING---NN	555
QY	589	YFYVDPREYEDLKWEFFRENLEFGKVLGSGAFGKVMNATAGISKTQVSIOQAVVKMLKE	648
DB	556	YVYIDPTQLPYDNKWEFFRDLRCFKGILGAGAFGVVENTAGLLKEDSRITVAVKMLKP	615
QY	649	KADSSREALMSCLKMTOLGSHENIVNLLGACTTUGPTIYLFICYCYGDLNLYRSKRE	708
DB	616	SAHSTEREALMSCLKVSLYGIHHKHNIVNLLGACTVGGTFTVITEYCCYGDLLNLYRRKD	675
QY	709	KFHRTWTEIFKEHNFSFYPTFOSHPSN---SMPGSRREVQIHPDSQI-----	752
DB	676	SF-----ICPKFEDNSEALYKNLLNTRDMGCEGMSXEYIDMKPAYSVYVPT	721
QY	753	--SGLHGNSFHSEDIETENQKRLEEEEDLVNLTPELFCFAYQAKGMEFLFKSCVHR	810
DB	722	KTDKRRSGSGDQV-----SVSIFEEDDL-ALDPTEDLINFYSQVQACMNLASKNCIHR	775
QY	811	DLAARNVLYTHCKVVKICDFGLARDIMSDSNVYVGNARLPVKWMAPESLFEGYITTKSD	870
DB	776	DUAARNILTLTHGRITKICDFGLARDIRDSNVYVGNARLPVKWMAPESEFHCYVTFESD	835
QY	871	VNSYGTLLWEIYFLSGVNPYGPPIVDANFYKLIQNGFKMDQFPYATEELIYIMQSCWAFDS	930
DB	836	VMSYGIILLWEIYFLSGSSPYPRIPVDYSKEYIKMIKGYRMKMSPECAPLEMYEIMRSCWSDP	895
QY	931	KRRPSFPNLTSLFGQLADAE	951

Db 896 LKRPTFKQIVQMVEQQLSDSK 916

RESULT 14

```

Q9I8NG PRELIMINARY; PRT; 977 AA.
AC Q9I8NG;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
Ems.
GN CSflr OR FMS.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
NCBI_TaxID=7955;
RX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=AB;
RX MEDLINE=20323166; PubMed=10862741;
RA Parichy D.M., Ransom D.G., Paw B., Zon L.I., Johnson S.L.;
RT "An orthologue of the kit-related gene fms is required for development
of neural crest-derived xanthophores and a subpopulation of adult
melanocytes in the zebrafish, Danio rerio."
RL Development 127:3031-3044(2000).
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
KINASES.
CC EMBL: AF240639; AAF76872.1; -.
DR HSSP: F11362; IFGK.
DR ZFIN: ZDB-GENE-001205-1; csflr.
DR InterPro: IPRO00719; Euk_pkinase

```

DR	InterPro: IPR003598;	Ig_C2.
DR	InterPro: IPR003006;	Ig_MHC.
DR	InterPro: IPR001824;	RTKkinaseIII.
DR	InterPro: IPR001245;	Tyr_kinase.
DR	Pfam: PF00047;	Ig; 4.
DR	Pfam: PF00065;	pkInase; 1.
DR	Prodom: PSD00001;	Euk_pkinase; 2.
DR	SMART: SM00408;	IGC2; 1.
DR	SMART: SM00219;	TyrcK; 1.
DR	PROSITE: PS00107;	PROTEIN_KINASE
DR	PROSITE: PS50011;	PROTEIN_KINASE
DR	PROSITE: PS00109;	PROTEIN_KINASE
DR	PROSITE: PS00240;	RECEPTOR_TYR_KIN
KW	ATP-binding; Glycoprotein; Immunoglobulin	
KW	Receptor; Transferrase; Transmembrane	
FT	VARIANT 238 238	N ->
FT	VARIANT 615 615	N ->
FT	SEQUENCE 977 AA;	110187 MW; C
SQ		

Query Match	23.2%;	Score 1223.5;	DB 13;	Length 977;
Best Local Similarity	31.6%;	Pred. No. 3.2e-92;		

QY	71	SSGTVVEAAAVEVDVSASITLQVLVDAPGNISCL-VWFKHSSLNCOPHFDLQNRGVVSMV	129
Db	27	SSAL-----AGTDVILESGSPQLVCGDGPVTLPLAKH-----KRYISKEVGKIRSR	78
QY	130	ILKMTETAQGEV-LFLIQSEANTYTLTFVSYRNT-LLY-----TLRRYPYFRKMNQDAL	182
Db	79	VEKTVDFGTGKYCVYVWNGDNLSSVHVFRDSRVLEFVSPSTLSR--YVRK-EGEDIL	135
QY	183	VCISESVPEPIVEWVLCDQSQESCKEESPAVYKKEKVL-----HELFGDIDICCCARNEIG	238
Db	136	LPCLLTDP-ATDFTFRMDNGSAAPYGMNITDPRKGVILRVHPGFNADYICCARIGCA	194
QY	239	RECTRLFTDLNQTQTTLPLQFLK-----VGPELWTRCAVHVHNGFGLTWELENKA	291
Db	195	EKVKIFSINIQRFP-PYVILKRVNKKVUGERLQISCTNNPNPFYNYTWTWTHSSRM	253
QY	292	LEEGNYFEMSTYSTNRTMIRILFAFVSSVARNDTGYTCTSSSKHPSSQALVT-----IVGRK	348



```

Db 254 LPKAE--EKSTMEGRUAISILT-IPFVQLSHSTGNITCTQNEAGANSSITQLLVWEEP 310
QY FI-----NATNSSDEYIDQYEEFCFSVRFKAYQIIRC-----TWTFSRKSPFCEQKG 396
Db 311 YIRLSPKSLKTHRGLSIEVSEGGDDVLGLVIEAYPLTSHKWTPTSHNASLP----- 365
QY 397 LDNGYSISKFCNHK-----HOPGEYIFHAENDDAOFTKMTNIRKPKOVL 442
Db 366 -----ENREFNHDRYEALLKRLNFAEETIGQYTLNKNMSKASITFDIKMTKPYAR 419
QY 443 AE-ASASQASCFSDGYPLPWTWKCKSDKSPNCEETE-----GVNRKANRKFVGQWVS 497
Db 420 VKWENVTLSCRSVGPAPSLWTLQCTGIRTPCPTENTDLOPIQTQVDFQKESFGAVGV 479
QY 498 SSTLNINSEAIKGFVLKVCAYNSLGTSETILLNSPGFPFPFQDNISFYATIG--VCLIFI 555
Db 480 ESVLTGPN--RMVTVCVAENLVGGSDTFSME-----VSDQIFTSAMCGSVAMVVL 531
QY 556 VLTLLICHYKQKQFRYESQLQWQVVTGSSDNEYFYDFREYEDLKWEPFRENLEFGKV 615
Db 532 GLLIFMIYKQKQPRYEIRKWIIEAT--NGNNYTFIDPTQLPYNEKWEFPRDKLKGKT 589
QY 616 LGSAGFKVMNATAYGISKTCVSTQVAVKMLKEKADSEREALMSELKMTQOLGSHENIV 675
Db 590 LGAGAFKVVVEATAYGLGEDNITRVAVKMLKASHPDEREALMSELKILSHLGOKNIV 649
QY 676 NLLGACFLSGPIYLIFCYCYGDLNLYLRSKREKF-----HRT 713
Db 650 NLLGACHTGGPVLVITYCYCHGDLNLFRLSKAENFLNFMVTIPNFPPEMTDYKNVSTERM 709
QY 714 W-----TEIFKEHNSFYPTFQSHPNSSMPGSRVQIHPDSDQISGLHNSFHSDEI 766
Db 710 FVRSDSGISSTCSHYLDMRPVTSRPTNSALDSSECC----- 747
QY 767 EYENOKLEEBEDLNVTFFEDLLCFATQVAKGMFLEFKSCVHRDLAARNVLVTHGKVKV 826
Db 748 -----EDSWPLDMDDLLRFSSQVAGGLDFLAAKNQIHRDVAARNVLLTNSRVAK 796
QY 827 ICDFLARDIMSDSNVYVGRNARLPVKWMAPELFEGLITIKSDVWSYGILLWEIFSLGV 886
Db 797 ICDFLARDIMSDSNVYVGRNARLPVKWMAPELFEGLITIKSDVWSYGILLWEIFSLGK 856
QY 887 NXPYGPVDANFYKLIQNGFKMDQPFYATEEIIYIMOSCAWFDKSRKPSFPNLTSLGCG 946
Db 857 SPYPNILVDSKFKYKMGQYQMSRPDFAPEPMYTIMKMCWNLDAAREPTFSKISQMIQRM 916
QY 947 LADAEAM-----YQNV-----DGRVSECPTYQNRPPF-----SREMDLGLLSP 986
Db 917 LGSETSEQDTQYKNIPTAEAEQOLESQDPVKHDESFETSCDQEEBDOPLMKP 971

```

## RESULT 15

```

Q9TDT7
ID Q9TDT7 PRELIMINARY; PRT: 948 AA.
AC Q9TDT7
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Stem cell factor receptor (Fragment).
GN C-KIT.
OS Trichosurus vulpecula (Brush-tailed possum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
OX NCBI_TaxID=9337;
RN [1]
RP Lawrence S.B., Greenwood P., Tisdall D.J., McNatty K.P., Fidler A.E.;
RT "Partial cDNA sequence of the c-kit homologue of brushtail possum
RT (Trichosurus vulpecula).";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF131209; AAF22141.1; -
DR HSSP; P11362; 1FGK.

```

```

DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR00306; Ig_MHC.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00047; Ig; 3.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_pkinase; 2.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Immunoglobulin domain; Receptor; Tyrosine-protein kinase.
FT NON_TER 948
SQ SEQUENCE 948 AA; 106323 MW; 708E0258FFFC4D07D CRC64;

```

Query Match 23.2%; Score 1221; DB 6; Length 948;

Best Local Similarity 33.2%; Pred. No. 4.9e-92;

Matches 319; Conservative 160; Mismatches 363; Indels 118; Gaps 32;

```

QY 67 LRPOSSGTVYEAATAVEVDVDSASITLQVLVDAPGNISCL-----WVFKHSSLNCQPHFDL 120
Db 24 LQSTSPANYSSPSI---TPARSQLTVNFSGSEIKLSCSDLHFVNMTFFENAHLLE----- 75
QY 121 QNRGVVSVILKMTETQAGEYLLFTQSEATNTYTLFTVSIKNTLLYTLRRPYFRKMNQD 180
Db 76 SSRTEETLTTNAKAQDTGRYCTCTNKMGLSSIIYVFKDPK--ILFLDLRLLYGN-EGSD 132
QY 181 ALVCISESVPEPIEVW--LCD-----SQGESCKEESPAVVKKEKVLHEL 224
Db 133 ALVNCPTVDPE-VTNFTLRLCDGKPLKDLTLIPDIQKGIITKN-----VKRSHNKIC--- 184
QY 225 FGTDTRCCARNELGRETRLTIDLNQTPOTTLPOL-----FLKVGEPMLIRCKAVHV 277
Db 185 ----PQSAAYOGGQKLSQDMFLKVRPAQ--SVPEVSLLOTWYLLREGFTFOATCMIKDV 239
QY 278 NHGFLTWELNKALEEGNYFEMSTYSTNRMTIRILFAFVSSVARNDTGYTSCSSK--- 334
Db 240 ASSVVSWMKIDNNRI---STHTQSRHSGDYAYERQDILTISPVRVNDSGVTCFANNTG 296
QY 335 HPSQSALVTYVGKGINA-TNSEDYEIDQEEFCFSVRFKAYQ-----TRCTWTFSR 387
Db 297 PANVTATLKVVVEKGFNIFPKMNTTIFINDGENIDLVWEYEAEPKPEHLQWIMNGTVD 356
QY 388 KSEPCQKGLDNGYISIKFCNHK-----OPGEYIFHAENDDAOFTKMTNIRKPKQ 440
Db 357 KWDDYTKGCE---STIRIISLHLNRLKGTGGPYTFSVSNDSVSTSVTFKVVYKTRPE 413
QY 441 VLAESASQA--SCFSQGYPLPSWTWKCKSDKSPNCEETIEGVMNRKANRV---FGQM 495
Db 414 ILTSDRLMGLLQCVAAAGFPPEPTIDWYFCPGTEQRCSTSVLP--MDVKTNSMMLPPFGKI 471
QY 496 VSSSTLNINSEAIKGFVLKVCAYNSLGTSETILLNSPGFPFPF-IDNISFYATIGVCLL- 553
Db 472 VVESTIDSSAFRYNGTVECKASNDVG-----KSSAFFNFAIKEQITSHLFTPLLG 523
QY 554 FIV-----VLTLLICHYKQKQFRYESQLQWQVVTGSSDNEYFYDFREYEDLKWEEP 606
Db 524 FVAAGLMCVIILTYIKYFQKPMYEVQWQVVEEING---NNYVYIDPTQLPYDHRKEEP 580
QY 607 RENLEFFGVLGSGAFKVMNATAYGISKTCVSTQVAVKMLKEKADSEREALMSELKMT 666
Db 581 RNRLSFGKTLGAGAGKVVVEATAYGLFKSDAAMTVAVKMLKPSAHLTEREALMSELKVL 640
QY 667 QLGSHENIVNLGACTLSGPIYLIFCYCYGDLNLYLRSKREKFHRTWTEIFEKH-----N 722
Db 641 YLGNHNMNIANLLGACTIGGPTLVITEYCYGDLNLFRLKRRKDSF-----ICKSHEDHAE 694
QY 723 PSFPTFQSHPNSSMPGSRV-VQIHPDSDQISGLHNSFHSF-----DEIYENQKRLLEE 778
Db 695 AALYKNLLQSKESCDGANEMYDMKPGVSYVPTKAEKRRRSARVGSYTERDVTTAIMEDD 754

```

